ிர்ப் புள்ளியியல்

(BIO STATISTICS)

S. S. நாராயணசாமி

தமிழ்நாட்டுப் பாடநூல் நிறுவன

த பா. நி. (க.வெ.) வரிசை எண்—344

உயிர்ப் புள்ளியியல்

ஆசிரியர்

எஸ். எஸ். நாராயணசாமி, எம்.எஸ்ஸி.

புள்ளியியல் துணைப் பேராசிரியர்,

சென்னை மருத்துவக் கல்லூரி,

தமிழ்நாட்டுப் பாடநூல் நிறுவனம்

First Edition—August, 1972

T.N.T.B.S. (C.P.) No. 344

BIOSTATISTICS

S. S. NARAYANASWAMI

Price Rs. 12-30

'Published by the Tamil Nadu Text Book Society under the Centrally Sponsored Scheme of Production of books and literature in regional languages at the University level, of the Government of India in the Ministry of Education and Social Welfare (Department of Culture), New Delhi.'

Printed by
BHARANI PRESS,
11, Venkatadri Naicken Street,
Kuyappettai, Madras-12.

அணிந்துரை

திரு. இரா. நெடுஞ்செழியன்,
(தமிழகக் கல்வி—உள்ளாட்சித் துறை அமைச்சர்)

தமிழைக் கல்லூரிக் கல்வி மொழியாக ஆக்கிப் பன்னிரண்டாண்டுகள் ஆகிவிட்டன. குறிப்பிட்ட சில கல்லூரிகளில் பி.ஏ. வகுப்பு மாணவர்கள் தங்கள் பாடங்கள் அனைத்தையும் தமிழிலேயே கற்று வந்தனர். 1968ஆம் ஆண்டின் தொடக்கத்தில் புகுமுக வகுப்பிலும் (P.U.C.), 1969ஆம் ஆண்டிலிருந்து பட்டப் படிப்பு வகுப்புகளிலும் அறிவியல் பாடங்களையும் தமிழிலேயே கற்பிக்க ஏற்பாடு செய்துள்ளோம். தமிழிலேயே கற்பிப்போம் என முன்வந்துள்ள கல்லூரி ஆசிரியர்களின் ஊக்கம், பிற பல துறைகளிலும் தொண்டு செய்வோர் இதற்கெனத் தந்த உழைப்பு, தங்கள் சிறப்புத் துறைகளில் நூல்கள் எழுதித் தர முன்வந்த நூலாசிரியர்கள் தொண்டுணர்ச்சி இவற்றின் காரணமாக இத் திட்டம் நம்மிடையே மகிழ்ச்சியும் மனநிறைவும் தரத்தக்க வகையில் நடைபெற்றுவருகிறது. இவ் வகையில், கல்லூரிப் பேராசிரியர்கள் கலை, அறிவியல் பாடங்களை மாணவர்க்குத் தமிழிலேயே பயிற்றுவிப்பதற்குத் தேவையான பயிற்சியைப் பெறுவதற்கு மதுரைப் பல்கலைக்கழகம் ஆண்டுதோறும் எடுத்துவரும் பெருமுயற்சியைக் குறிப்பிட்டுச் சொல்லவேண்டும்.

பல துறைகளில் பணிபுரியும் பேராசிரியர்கள் எத்தனையோ நெருக்கடிகளுக்கிடையே குறுகிய காலத்தில் அரிய முறையில் நூல்கள் எழுதித் தந்துள்ளனர்.

வரலாறு, அரசியல், உளவியல், பொருளாதாரம், தத்துவம், புவியியல், புவியமைப்பியல், மனையியல், கணிதம், இயற்பியல் வேதியியல், உயிரியல், வானியல், புள்ளியியல், விலங்கியல், தாவரவியல், பொறியியல் ஆகிய எல்லாத் துறைகளிலும் தனி நூல்கள் மொழிபெயர்ப்பு நூல்கள் என்ற இரு வகையிலும் தமிழ் நாட்டுப் பாடநூல் நிறுவனம் வெளியிட்டுவருகிறது.

இவற்றுள் ஒன்றான 'உயிர்ப் புள்ளியியல்' என்ற இந் நூல் தமிழ்நாட்டுப் பாடநூல் நிறுவனத்தின் **344**ஆவது வெளியீடாகும். இதுவரை **379** நூல்கள் வெளிவந்துள்ளன. இந் நூல் மைய அரசு கல்வி, சமூக நல அமைச்சகத்தின் மாநில மொழியில் பல்கலைக்கழக நூல்கள் வெளியிடும் திட்டத்தின்கீழ் வெளியிடப்படுகிறது.

உழைப்பின் வாரா உறுதிகள் இல்லை; ஆதலின், உழைத்து வெற்றி காண்போம். தமிழைப் பயிலும் மாணவர்கள் உலக மாணவர்களிடையே சிறந்த இடம் பெறவேண்டும். அதுவே தமிழன்னையின் குறிக்கோளுமாகும். தமிழ்நாட்டுப் பல்கலைக் கழகங்களின் பல்வகை உதவிகளுக்கும் ஒத்துழைப்புக்கும் நம் மனம் கலந்த நன்றி உரித்தாகுக.

இரா. நெடுஞ்செழியன்

பொருளடக்கம்

1. தோற்றுவாய்

தனித்துறைப் பயிற்சி தற்கால நியதியாக அமைந்துள்ளது. ஆய்வின் தன்மை அதிகமாக அதிகமாகத் தனித்துறைப் பயிற்சிக்கான களங்களும் அதிகமாகின்றன. செயற்கள எல்லை வரையறுக்கப்பட்டுக் குறுகியதாய் அமைந்தாலும், செயலின் தன்மையில் மிக்க கவனமும் தீவிரத்தன்மையும் கொள்ளவேண்டி உள்ளது. புதுக்கருத்துகளைச் சேர்க்கவேண்டியவர்களாகவும், பழைய கருத்துகளைச் சோதனைக்குள்ளாக்கவோ அல்லது திருத்தி அமைக்கவோ வேண்டியவர்களாகவும் உள்ளோம். பெற்ற அறிவினை, ஒத்த துறையில் முயற்சி செய்து வரும் சக ஊழியர்களுடனும், உள்நாட்டிலும் வெளி நாட்டிலுமுள்ள அத்துறை மற்றும் அத்துறையைச் சார்ந்த பிரிவில் செயல்படும் வல்லுநர்களுடனும் கருத்துப் பரிமாற்றத் தேவையை உணர்கிறோம். உயிரியல் துறை மட்டும் இதற்கு விதிவிலக்காக அமையமுடியாது. புதிய கருத்துகளைப் பெறவும், பெற்ற கருத்தினை மற்றவர்கள் சந்தேகத்திற்கிடமின்றி ஏற்றுக்கொள்ளும் தன்மையில் கூறவும் புள்ளியியலின் உதவியை நாட வேண்டியுள்ளது. மற்றவர்களால் ஏற்றுக்கொள்ளும்படி அமையும் கருத்துகள் ஆய்வறிவு சார்ந்த கருத்துகளாக அமைய வேண்டியனவாயும், ஆய்வறிவு சார்ந்த கருத்துகள் புள்ளியியல் முறைப்படி அணுகியனவாயும் அமையவேண்டி உள்ளன.

புள்ளியியல் உத்திகள் அறிவியல் முறையின் ஒரு நோக்கமாகும். அறிவியல் முறைக்குச் சுருக்கமான சொற்பொருள் விளக்கம் தருவது எளிதன்று. ஆராய்ச்சிக்குரிய செய்தியைப் புரிந்து கொள்வதற்காக வாதமுறை (logic), மற்றும் புறநோக்கு (objectivity) முறைகளைப் பயன்படுத்துதல் அறிவியல் முறை எனச் சொல்லலாம். முன்பு கண்டறிந்த செய்திகளின் தன்மையைச் சோதனை செய்து, சோதனைக்குள்ளாக்கப்பட வேண்டிய கருதுகோள்களை (hypotheses) அமைப்பதே அறிவியல் முறையின் முக்கிய நோக்கமாகும். கருதுகோள்களை அமைக்கும் திறமை

தனிப்பட்டவர்களின் அனுபவத்தைச் சார்ந்ததாகும். சிறந்த கருதுகோள்களை அமைப்பதற்கான நியதிகள் கொடுக்கப்பட்டிருந்தாலும், குறிப்பிட்ட ஆய்விற்குப் பொருத்தமான கருதுகோள்களை அமைப்பது, கருதுகோளை அமைப்பவர்தம் கூர்மதியையும் (genius) உள்ளுணர்வையும் (intution) சார்ந்தமையும். கருதுகோட் சோதனைகளுக்கான (testing of hypotheses) அளவுகளைச் சேகரிப்பதிலும், கருத்தாய்வுத் தொகுதியிலிருந்து (theory) முன்கூட்டி உணரத்தக்க கருத்துகளைக் கூறுவதற்கான அளவைகளின் ஒப்பீட்டு முறைகளைக் கொடுப்பதிலும் புள்ளியியல் துறை, அறிவியல் துறைக்குத் துணைசெய்கிறது. புள்ளியியல் துறையைத் தவிர்த்து, வேறெந்தத் துறையிலும் புள்ளியியல் தனித்து நின்று புதிய அறிவியல் கோட்பாடுகளை உண்டாக்க முடியாது. அமைக்கப்பட்ட கோட்பாடுகளைச் சோதனைக்குள்ளாக்கப் புள்ளியியல் உதவுமேயன்றிப் புதிய கருத்துகளை உண்டாக்கப் புள்ளியியல் உதவுவதில்லை.

அறிவியல் துறையில் புள்ளியியல் முறைகள் மூன்று வழிகளில் பயன்படுத்தப்படுவதைக் குறிப்பிடலாம். விரித்துரைத்தல் (description), ஆய்தல் (analysis), முன் மதிப்பிட்டுரைத்தல் (prediction) ஆகிய செயல்களைப் புள்ளியியல் முறைகள் செய்கின்றன. விரித்துரைத்தலாவது, கொடுக்கப்பட்ட செய்திகளைச் சுருக்கி அமைத்துத் தகுதியான புள்ளியியல் அளவைகளால் Statistical constant) வருணிப்பதைக் குறிக்கும். கொடுக்கப்பட்ட செய்திகளை வருணிப்பதுடன், இரு முழுமைத் தொகுதிகளை (population) ஒப்பிடவும் இவ் வளவைகள் உபயோகப்படுகின்றன. விரித்துரைத்தலாவது பெற்றறி (deductive) முறையாகவோ அல்லது பட்டறி (inductive) முறையாகவோ இருக்கலாம். முழுமைத் தொகுதி அளவுகள் முழுதும் தெரிந்தவிடத்துச் செய்யும் ஆய்வு முறைகளையோ, அல்லது முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட மாதிரி (Sample) அளவுகளிலிருந்து செய்யப்படும் ஆய்வு முறைகளையோ இவை குறிக்கின்றன. மாதிரி அளவுகளிலிருந்து முழுமைத் தொகுதிக்கான பண்புகளைக் கணக்கிடுவதே பெரும்பாலான ஆய்வு முறைகளின் நோக்கமாகும். கொடுக்கப்பட்ட அளவுகள் ஒரு குறிப்பிட்ட முழுமைத் தொகுதியைச் சேர்ந்ததா அல்லது எடுக்கப்பட்ட இரு மாதிரிகளின் ஓர் அளவைக்கான வித்தியாசம் குறிப்பிடும்படி (significant difference) உள்ளதா எனும் கருதுகோட் சோதனைகளில் புள்ளியியல் முறைகள் பெரிதும் பயன்படுகின்றன. மாதிரி அளவைகளிலிருந்து (statistic) முழுமைத் தொகுதி அளவைகளை (parameter) மதிப்பீடு (estimation) செய்வதற்கும் புள்ளியியல் முறைகள் உபயோகப்படுகின்றன. மாதிரி உறுப்புகள் யாவும் ஒரு குறிப்பிட்ட முழுமைத்

தொகுதியிலிருந்து வந்தன என்ற அடிப்படையில் இம் மதிப்பீடுகள் செய்யப்படுகின்றன. இம் மதிப்பீடுகளின் உதவியால் முழுமைத் தொகுதியின் உருவமைப்பு (model) முழுவதுமாக வருணிக்கப்படுகிறது. இவற்றின் உதவியாலேயே புள்ளியியலின் மூன்றுவது செயலான வருவதுரைத்தல் செம்மைப் படுத்தப் படுகிறது. எனவே,

'புள்ளியியல் இல்லாத அறிவியல் பயன்தராது
அறிவியல் இல்லாத புள்ளியியல் வேரற்றது'

என்பதான கூற்று மிகப் பொருத்தமாகக் கூறப்பட்டதெனக் கொள்வது மிகையன்று.

அளவுசார் உயிரியல் (Quantitative Biology) பிரிவில் புள்ளியியல் உபயோகம்பற்றி வாதிடும் நிலை மாறிவிட்டது. இதில் ஏற்படும் வாதம், உயிரியல் வல்லுநர்களால் எளிதாகப் புரிந்து கொள்ளும்படி புள்ளியியல் முறைகள் அமையவில்லை என்று எண்ணுவதாலோ அல்லது தேவையில்லாத இடங்களில் கூடப் புரிவதற்கு மிகக் கடினமான சில நுட்பமான புள்ளியியல் செயல் முறைகளை மிகைப்படுத்திக் கூறுவதாலும்தான் அமைந்துள்ளது. புள்ளியியல், உயிரியல் பிரிவிற்குத் தேவையா என்பதைவிட அதன் தேவையை உணர்ந்து அதனை அணுகும் முறையில் உள்ள தன்மையையே இவ்வாதம் கூறுகிறது. எல்லா உயிரியல் வல்லுநர்களும் புள்ளியியல் முறைகளை வெளிப்படையாக உபயோகப் படுத்தாவிட்டாலும், ஏதாவது ஒரு வழியில் உள்ளுணர்விலாவது பயன்படுத்துவதைக் காணமுடிகிறது. மனிதர்களையும், வாலில்லாக் குரங்கினையும் வேறுபடுத்திக் காட்டுவது மிக எளிது. கொடுக்கப்பட்ட இரட்டையர்கள் ஒற்றைக் கருவிலிருந்து தோன்றியவர்களா அல்லது இரட்டைக் கருவிலிருந்து தோன்றியவர்களா என்பதை உடலின் புறத்தோற்றத்தைக் கொண்டு கூறுவது சிக்கலாக அமைந்துள்ளது. ஆதலின் உயிரியல் பிரிவில் புள்ளியியலின் முக்கியத்துவத்தை ஏற்றுக்கொள்வது தொடர்ந்து செல்வதற்குச் சாதகமாக அமைகிறது.

மற்ற எல்லா அறிவியல் துறைகளைப் போலவே உயிரியலும் கண்டறிதல் மற்றும் அளவுகளின்பால் சார்ந்ததாகும். உயிர்ப் பொருள்கள் அனைத்தும் ஒன்றிலிருந்து ஒன்று மாறுபட்டு அமைந்துள்ளன. கொடுக்கப்பட்ட அளவுகளில் உள்ள மாற்றங்களைக் கணிக்கும் முறையில் நம்மைத் தவறான கருத்துகளுக்கு இட்டுச் செல்லாமல், அளவுகளின் தன்மையைக் கூறப் புள்ளியியல் முறைகள் ஓர் இன்றியமையாத உத்தியாக அமைந்துள்ளது.

உயிருடன் இருக்கும் அல்லது உயிருடன் இருந்த பொருள்களின் அளவுகளைப் புள்ளியியல் முறைகளைக் கொண்டு படிக்கும் துறையே உயிர்ப் புள்ளியியல் (Biostatistics) எனப்படும். விலங்கு மற்றும் தாவர இனம் ஒன்றின் அளவுகளை வைத்துக்கொண்டு உயிரியல் மற்றும் புள்ளியியல் வல்லுநர்கள் சோதனை செய்யும் சக்தியை ஒருங்கிணைத்து ஆராயும் பிரிவே உயிர்ப் புள்ளியியல் எனவும் கூறலாம்.

உயிர்ப் புள்ளியியல் எனும் சொல் பத்தொன்பதாம் நூற்றூண்டின் இறுதியிலிருந்து கையாளப்பட்டு வருகிறது. ஒரு வகையில் உயிரியல் அளவுகளை விதிமுறை தவறு நுட்பமான முறைகளில் கணக்கிடுவதையும், மற்றொரு வகையில் வரையறுக்கப்பட்ட கணக்கியல் பகுப்பு முறைகளைக் கையாளுவதையும் இது குறித்தது. 'உயிர் அடிப்படையான புலன்களால் உணரத்தக்க அளவுகளுக்குரிய பண்புகளைப் படிப்பதே' உயிர்ப் புள்ளியியலின் முக்கிய நோக்கமாகும். காலப்போக்கில் பண்பின் அடிப்படையில் குறிக்கப்படும் உயிரியல் தன்மையை உயிரியல் பிரிவாகவும், அளவுகளின் அடிப்படையில் குறிக்கப்படும் உயிரியல் தன்மையை உயிர்ப் புள்ளியியல் பிரிவாகவும் கொண்டால் வியப்படைவதற்கில்லை. உயிரியலைப் பொது அளவைக்குள்ளாக்குவதே உயிர்ப் புள்ளியியலின் முக்கிய வேலையாகும். உயிரியல் நிகழ்ச்சிகளைப் பொது அளவைக்குள்ளாக்காமல் அவற்றைப் படிப்பதற்குத் தேவையான முழு அறிவையும் பெற முடியாது என்பதே அதன் அடிப்படை நியதியாகும். அளவு குறிப்பிடாமல் விவரிக்கப்படும் ஒரு பொருளின் தன்மை முற்றுப் பெற்ற கருத்தாக அமையாததோடு அடிப்படை முக்கியத்துவத்தையும் இழந்து விடுவதை நம்மால் உணரமுடிகிறது. நிகழக் கூடிய தவற்றினைக் கணக்கிட முடியாதபடி அமைக்கப்படும் சோதனை முறைகள் உண்மையான அல்லது தவறான முடிவுகளைக் கூறுவதில் சமவாய்ப்பினை உடையதாய் அமைந்துவிடுகிறது.

நமக்கு எண்களிடத்தில் எத்தகைய வெறுப்பு இருப்பினும் நமது செயல்களையும், திட்டங்களையும் புள்ளியியல் முறைப்படி திட்டமிட்டுச் செயல்படுத்தி, மதிப்பிட வேண்டியவொரு தவிர்க்க முடியாத சூழலில் உள்ளோம். மூல அளவுகளுக்கான பண்புகளைச் செம்மையாகக் கூறவும், அவற்றில் உள்ள குறைபாடுகளின் தன்மையைத் தெளிவாக்கவும் இத் துறை பயன்படுகிறது. அறிவுத் துறையில் ஏற்பட்டுள்ள புதுச் சூழலில் உயிரியல் வல்லுநர்கள் புள்ளியியல் முறைகளைப் புரிந்துகொள்வது அவர்தம் செயல்களைச் சரியாகத் திட்டமிடவும், முறைப்படி செயல்படுத்தவும், செய்திகளைத் தொகுக்கவும், தொகுத்தவற்றைப்

பகுக்கவும், பகுத்தவற்றிலிருந்து பெறப்படும் கருத்துகளைத் தெளிவாக அறுதியிட்டுக் கூறவும், புள்ளியியல் வல்லுநர்களால் கூறப்படும் கடினமான புள்ளியியல் முறைகளைப் புரிந்துகொள்ள வும் இன்றியமையாததாகும். காலம், பொருள், சக்தி முதலியன வீணுகாமல் செயல்முறைகளை ஒழுங்குபடுத்திச் செயல்புரிய இத்துறை பெரிதும் பயன்படுகிறது. குறைந்த செய்திகளிலிருந்து நிறைந்த கருத்துகளைப் பெற இத்துறையைப்போல் வேறெந்தத் துறையும் இல்லை எனச் சொன்னால் மிகையன்று.

புள்ளியியல் உதவி இல்லாமல் உயிரியல் பிரிவில் தோன்றும் சிக்கல்களை எளிதாக்க முடியாதென்பதைத் தக்க சான்றுகளுடன் காண உள்ளோம். புள்ளியியல் முறைகளின் தற்கால வளர்ச்சியினை நன்கு புரிந்துகொண்டாலன்றி உயிர்ப்புள்ளியியலை உயிரியலார் செம்மையாகப் பயன்படுத்த முடியாத நிலை உறுதிப்படினும் உயிர்ப் புள்ளியியலின் அடிப்படைக் கருத்துகளைக் கூறுவதே நமது நோக்கமாகும். ஒவ்வொரு உயிரியல் வல்லுநரும் தேர்ந்த புள்ளியியல் வல்லுநராகவும் இருக்க வேண்டுமென்று எதிர்பார்ப்பது தேவைக்கு அதிகமான அல்லது நடைமுறைக்கு ஒவ்வாத ஒன்றுக இருந்தாலும், உயிரியலில் பெறப்படும் கருத்துகள் யாவும் உயிர்ப் புள்ளியியல் முறைப்படி தெளிவு படுத்தப்பட்டுச் சந்தேகத்திற்கிடமின்றி எல்லோராலும் ஒத்துக் கொள்ளப்படுவதை மறுக்கமுடியாது. உயிரியல் பிரிவில் ஏற்படும் ஒவ்வொரு முடிவின் தன்மையும் உயிரியலார் உயிர்ப் புள்ளியியலின் அடிப்படைச் செயல் முறைகளை ஆயுதமாகப் பயன்படுத்தவேண்டிய நிலையில் உள்ளது. உறுதியான, சிறப்பான மற்றும் நம்பும்தன்மைத்தான கருத்துகளைக் கூற உயிர்ப் புள்ளியியல் பெரிதும் பயன்படுகிறது. நமது கருத்துகளைக் கூற இந்த ஒரு முறைதான் உதவியாய் அமையும் என்று கூற முடியாவிட்டாலும், நமக்குள்ள எல்லா முறைகளிலும் இதுவே தலைசிறந்ததென்னின் தவருகாது. உயிர்ப் புள்ளியியல் முறைப்படி அணுகிக் கூறப்பட்ட கருத்துகள் குறைவான தவறுகள் உள்ளனவாய் அமைகின்றன. வேறெந்த முறைக்கும் இல்லாத, தனக்கே உரித்தான 'செயலில் ஏற்படும் தவறுகளை முன்கூட்டியே அளந்து கூறும் தன்மைத்தாய் அமைந்துள்ள' இதனை மறுக்க முடியாமல் கையாள வேண்டிய நிலையில் உள்ளோம்.

உயிர்ப் புள்ளியியலை உயிரியல் பிரிவில் உறுதியாக ஏற்றுக் கொள்ளப்பட்ட நிலையில் உயிர்ப் புள்ளியியலான கட்டடம் கால்டன் பெருமகனால் (Sir Francis Galton) அடிக்கல் நாட்டப்பட்டு பியர்சன் (E. S. Pearson) மற்றும் ஃபிஷர் (R. A. Fisher) ஆகிய வல்லுநர்களால் கட்டடமாக்கப்பட்டுள்ளதை நினைவு கூர்வது நமது கடமையாகும்.

2. நிகழ்வெண் பரவல்

(Frequency Distribution)

களத்தில் சேர்க்கப்பட்ட செய்திகளே புள்ளிவிவர ஆய்விற்கான அடிப்படைப் பொருள்களாக அமைகின்றன. உயிர்ப் புள்ளியியலுக்குத் தேவையான செய்திகள் மக்கள் தொகைக் கணக்கீட்டில் சேர்க்கப்படும் செய்திகளிலிருந்தும், பிறப்பு இறப்புக் குறிப்புகளிலிருந்தும், களத்திலும் சோதனைச்சாலைகளிலும் செய்யப்படும் சோதனைகளில் கிடைக்கும் அளவுகளிலிருந்தும் கிடைக்கின்றன. நமது தேவைக்கேற்ப இவற்றின் தன்மையும், அளவும் அமையும். தொகுக்கப்பட்ட விவரங்கள் முறையாக அமைக்கப்படாமலிருக்குமாயின் தெளிவான கருத்துகளை எளிதாகப் பெறமுடியாது. இவற்றை முறைப்படுத்தி அமைப்பதன் மூலம் தெளிவான கருத்துகளைப் பெறுவதோடு இச் செய்திகளைப் புள்ளிவிவர ஆய்விற்கான எளிதான பொருள்களாகவும் அமைக்க முடிகிறது.

வகைப்படுத்துதல் பண்பின் அடிப்படையினதாயும், அளவின் அடிப்படையினதாயும் இருக்கலாம். நம் நாட்டின் மக்கள் தொகையை ஆண், பெண் என்றும் முகம்மதியர், கிருத்துவர், இந்துக்கள் என்றும் படித்தவர்கள், படிக்காதவர்கள் என்றும் பிரிப்பது பண்பின் பகுப்புமுறைக்குக் காட்டாகவும், மக்கள் தொகையை வயதின் அடிப்படையில் பிரித்து ஒவ்வொரு வயதில் (அல்லது வயதுப் பிரிவில்) உள்ளவர்களின் எண்ணிக்கையைக் கொடுப்பது, அளவின் அடிப்படையில் பிரிப்பதற்குக் காட்டாகவும் சொல்லலாம். பண்பின் அடிப்படையில் பிரிப்பதையும் அதன் அளவை முறைகளையும் 'பண்புகளின் இயைபு' என்ற தலைப்பில் விரிவாகக் காண உள்ளோம். இங்கு, அளவின் அடிப்படையில் பகுப்பதையும் பகுத்த செய்திகளை வரைபடங்களாக அமைக்கும் முறைகளையும் தெளிவாக்குவோம்.

ஒரு பொருளிலிருந்து மற்றொரு பொருளுக்கு மாறுபடுகின்ற அளக்கும் தன்மைத்தான (Measurable characteristic) குணங்களை மாறிகள் (variables) என்கிறோம். பிறப்பு, இறப்பு விகிதங்கள், மரத்தில் உள்ள இலைகள், பூக்களின் இதழ்கள், நமது எடை, உயரம் முதலியன மாறிகளுக்குக் காட்டாகச் சொல்லலாம். குறிப்பிட்ட எல்லைக்குள் எந்த அளவினையும் ஏற்கும் மாறிகளைத் தொடர்மாறிகள் (Continuous variable) எனவும், முழு எண்களையே அளவுகளாகக் கொள்வனவற்றைத் தனித்த மாறிகள் (discrete variable) எனவும் கொள்கிறோம். குடும்பத்திலுள்ள ஆண், பெண் ஆகியவர்களின் எண்ணிக்கை, ஒரு செடியில் உள்ள பூக்களின் எண்ணிக்கை ஆகியவற்றைத் தனித்த மாறிக்கும், மழையின் அளவு, வெப்பத்தின் அளவு முதலியன தொடர் மாறிக்கும் காட்டாகச் சொல்லலாம். உயரம், எடை முதலியன தொடர் மாறிகளானாலும், நம்மால் முடிந்த அளவிற்குச் செம்மையாக அளந்து 70·5 கி.கி., 71·8 கி.கி., 63·8 கி.கி., 5அடி 6 அங்., 6 அடி 3அங், 4அடி 8 அங் என முழு எண்களாகவே குறிக்கிறோம். இது நமது தேவையையும், அளக்கும் கருவிகளின் நுட்பத்தையும் பொருத்து அமைகிறது.

தொகுக்கப்பட்ட செய்திகள் முறைப்படி அமைவதற்குமுன் தெளிவில்லாததாகவோ அல்லது சந்தேகத்திற்கிடமானதாகவோ அமைந்திருக்கலாம். தெளிவானதும், உறுதியானதுமான கருத்துகளைப் பெறச் செய்திகளை ஒழுங்குபடுத்தி அமைக்க வேண்டி உள்ளது. உயிர்ப்புள்ளியியலில் 60 மாணவர்கள் பெற்ற மதிப்பெண்களைக் கவனிப்போம். (மொத்த மதிப்பெண் 50)

22	34	35	27	17	33	17	27	22	27	03	18
10	33	29	31	25	25	26	13	03	19	15	16
21	26	42	16	28	35	18	24	21	24	36	33
18	09	26	18	28	15	10	10	26	38	12	25
47	29	25	29	13	36	38	02	31	14	00	33

மாணவர்களின் அகர வரிசைப்படி அமைந்த இவ்வெண்கள் பகுக்கப்படாத ஒன்றாகும்.

இத்தொகுதியை இருபிரிவுகளாகப் பிரித்து ஒரு குறிப்பிட்ட மதிப்பெண்ணுக்கு அதிகமாக எத்தனை பேர், அவ்வெண்ணுக்குக் குறைவாக எத்தனை பேர் என அமைப்பது பகுப்பின் முதற்படியாகும். இம்முறைப்படி 24-ம் அதற்குக் குறைவாக மதிப்பெண்கள் பெற்றவர்கள் 29 மாணவர்களாகவும், 24-க்கு அதிகமான மதிப்பெண் பெற்றவர்கள் 31 மாணவர்களாகவும் அமைகிறார்கள்.

கொடுக்கப்பட்ட எண்களை ஏறு வரிசையிலோ (ascending order) அல்லது இறங்கு வரிசையிலோ (descending order) அமைப்பது பகுப்பின் இரண்டாம் நிலை எனக் கொள்ளலாம். மதிப்பெண்களின் இறங்கு வரிசையை அடியிற்காணும் பட்டியல் காட்டுகிறது.

00	10	13	16	18	22	25	26	28	31	33	36
02	10	14	17	18	22	25	26	28	31	34	38
03	10	15	17	19	24	25	27	29	33	35	38
03	12	15	18	21	24	26	27	29	33	35	42
09	13	16	18	21	25	26	27	29	33	36	47

இப் பகுப்புமுறை (வரிசைப் படுத்துதல்—array) மேற்கண்ட முறையைக் காட்டிலும் (இரண்டு பிரிவுகளாகப் பிரித்தல்) தெளிவானதாகவும், அதனைப் பெறுவதற்குத் துணை செய்வதாகவும் அமைந்துள்ளது.

ஒவ்வொரு அளவும் எவ்வளவு முறை அமையப் பெற்றிருக்கிறது என்பதை அடியிற்காணும் பட்டியல் காட்டுகிறது.

மதிப்பெண்	மாணவர்கள் எண்ணிக்கை	மதிப்பெண்	மாணவர்கள் எண்ணிக்கை	மதிப்பெண்	மாணவர்கள் எண்ணிக்கை	மதிப்பெண்	மாணவர்கள் எண்ணிக்கை
0	1	14	1	22	2	31	2
2	1	15	2	24	2	33	4
3	2	16	2	25	4	34	1
9	1	17	2	26	4	35	2
10	3	18	4	27	2	36	2
12	1	19	1	28	3	38	2
8	2	21	2	29	2	42	1
						47	1

மேற்கண்ட பகுப்பு முறைகளைக் காட்டிலும், இப் பட்டியல் அளவுகளை முறைப்படி அமைப்பதிலும் கருத்துகளைத் தெளிவு படுத்துவதிலும் ஒரு சிறப்பினை ஏற்படுத்தி உள்ளது. இன்னும் சிறப்பாக அமைந்து புள்ளியியல் ஆய்விற்கான களமாக அமைவதிலும், கருத்துகளைக் கணித முறையில் அளவிடுவதற்கு ஏதுவாகவும் உள்ளது நிகழ்வெண் பரவல் எனும் கீழ்க்காணும் அமைப்பாகும்.

நிகழ்வெண் பட்டியல்
(Frequency table)

இடைவெளிப்பிரிவு (Class Interval)	நிகழ்வெண் (Frequency)
0 – 5	4
5—10	1
10—15	
15—20	11
20—25	6
25—30	16
30—35	7
35—40	6
40—45	1
45—50	1
மொத்தம்	60

தொகுக்கப்பட்ட செய்திகளைச் சரியான பிரிவுகளில் குறிப்பிடுவதே நிகழ்வெண் பரவலாகும். இது ஒரு காரணப் பெயர். கொடுக்கப்பட்ட மூல அளவுகளைக் குறிப்பிட்ட பிரிவுகளுக்குள் காட்டும் பட்டியலாக அமைகிறது. விவரங்களின் மூலத் தன்மையில் சிறு இழப்பினாலேயே இத்தகைய சுருக்கம் அமைகிறது. நம்மால் அமைக்கப்படும் பிரிவுகளினால் ஏற்படும் இழப்பு நமக்குக் கிடைத்துள்ள தெளிவினைக் காட்டிலும் குறிப்பிடத் தக்கபடி இல்லாத நிலையில் சுருக்கம் அமைக்கப்பட வேண்டும்.

மேற்சொன்ன மாதிரியில் 60 மாணவர்களது மதிப்பெண்களும் 5-10, 10-15, ... 45-50, எனப் பத்துப் பிரிவுகளாகப் பிரிக்கப்பட்டுள்ளன. இவை ஒவ்வொன்றையும் பிரிவு இடைவெளி (class interval) எனவும், பிரிவு இடைவெளியின் எல்லை எண்களை, பிரிவுக் கீழ் எல்லை (lower limit), பிரிவு மேல் எல்லை (upper limit) எனவும் சொல்கிறோம். இவ்விரு எல்லைகட்கும் உள்ள வித்தியாசத்தை (difference) பிரிவின் அளவு (size) எனவும், பிரிவு எல்லைகளின் மைய எண்ணை, பிரிவெண் (mid point) எனவும் குறிப்பிடுகிறோம். இப் பிரிவுகளில் அமையும் அளவுகளின்

மொத்த எண்ணிக்கையே பிரிவு நிகழ்வெண் (class frequency) எனக் கொள்ளப்படும். 15—20 என்ற அமைப்பில், 15 என்பது பிரிவுக் கீழ்எல்லையாகவும், 20 என்பது பிரிவு மேல்எல்லையாகவும், இவ்விரு எல்லைகளின் வித்தியாசமாகிய 5 பிரிவு அளவாகவும், $17\cdot5\left(=\frac{20+15}{2}\right)$ என்பது பிரிவெண்ணாகவும் அமைகின்றன. இப்பிரிவின் நிகழ்வெண் 11 ஆகும். இந் நிகழ்வெண், எடுத்துக் கொள்ளப்பட்ட 60 மாணவர்களில் 11 மாணவர்களின் மதிப் பெண்கள் 15-20 எனும் பிரிவு இடைவெளியில் அமைந்துள்ளதைக் குறிப்பிடுகிறது.

பிரிவு இடைவெளிகளைச் சம அளவுடைய இடைவெளிப் பிரிவுகளாகவும் (equal class intervals), சம அளவில்லாத இடை வெளிப் பிரிவுகளாகவும் (unequal class intervals), நுழைவு தவிர்க்கின்ற இடைவெளிப் பிரிவுகளாகவும் (exclusive class Intervals), உள்ளடக்கிய இடைவெளிகளாகவும் (inclusive class intervals) அமைத்துக் கொள்ளலாம்.

0-5., 5-10., ... 45-50., என்ற மேற்கண்ட பிரிவு இடைவெளிகள் நுழைவு தவிர்க்கின்ற சம அளவு இடைவெளிப் பிரிவுகள் எனச் சொல்லப்படும். மாறியின் அளவு, பிரிவு மேல் எல்லையோடு ஒத்த எண்ணாக அமையும்போது, அவ் வெண்ணை எப் பிரிவில் சேர்ப்பது என்ற சந்தேகம் எழுகிறது இவ்விடத்தில் மேற்சொன்ன இடைவெளிகளை நுழைவு தவிர்க்கின்ற இடை வெளிப் பிரிவுகளாகக் கொண்டு, மாறியின் அளவு, பிரிவின் மேல் எல்லைக்குச் சமமாக இருக்கும்போது, அவ்வெண் அப் பிரிவில் நுழைவதைத் தவிர்த்து அடுத்த பிரிவில் சேர்த்து விடுகிறோம். இம்முறையைத் தொடர்ந்து செய்வதால் இதனால் ஏற்படும் பிழை குறிப்பிடத் தக்கதாய் அமையாது. 0-5., 5-10., ..., 45-50., எனக் கொண்ட பிரிவு இடைவெளிகளில் ஒவ்வொரு இடைவெளியின் கீழ் எல்லைக்குச் சமமான மதிப்புகளை அப் பிரிவில் சேர்த்தும் மேல் எல்லைக்குச் சமமான எண்களை அடுத்த பிரிவில் சேர்த்தும் நிகழ்வெண்பரவலை அமைக்கிறோம். 25 என்ற மதிப்பெண் பெற்ற நான்கு மாணவர்களும் 20-25 என்ற பிரிவைச் சாராமல் 25-30 என்ற பிரிவைச் சார்ந்தவர்களாகக் கொள்வது இம் முறைப்படி (exclusive method) பிரிப்பதற்குச் சமமாகும்.

உள்ளடக்கிய இடைவெளிகள் இச் சந்தேகத்திற்கு இடமின்றி அமைகின்றன. இம் முறைப்படி அமைக்கப்படும் பிரிவு இடை வெளிகள் 0-4., 5-9., 10-14., ... எனத் தெளிவாக அமையும். மாறிகளின் அளவுகள் எந்த இடைவெளியைச் சார்ந்தது என்பது

தெளிவாக உள்ளது. இவ் இடைவெளிப் பிரிவுகளில் கீழ், மேல் எல்லைகளுக்குச் சமமான அளவுகளும் அப் பிரிவுகளிலேயே சேர்க்கப் படுவதால் இவை உள்ளடக்கிய இடைவெளிப் பிரிவுகள் (Inclusive class interval) எனப்படும். தனித்த மாறிகளுக்கு இப் பிரிவு இடைவெளிகள் சிறந்த செயல் முறை அமைப்பாக உள்ளது. ஆனால், தொடர்மாறிகளுக்கு இவ் வமைப்பைக் கொள்வது முடியாத ஒன்றாகும். மாறியின் அளவுகள் தொடர்ந்து பிளவின்றி அமைவதால் இரு இடைவெளிகளுக்கிடையில் ஒரு பிளவை (gap) விட்டு இடைவெளிப் பிரிவுகள் அமைப்பது தொடர் மாறிகளுக்கு ஒவ்வாததாக அமைகிறது. சிரமமாகப் படினும் தொடர்மாறிகளுக்கு நுழைவு தவிர்க்கும் இடைவெளிப் பிரிவு களை அமைப்பதே முறையாகும்.

உள்ளடக்கிய இடைவெளிகள் இரண்டிற்கிடையில் உள்ள பிளவினை (gap) இரண்டாகப் பகிர்ந்து, பிரிவின் கீழ் எல்லை யிலிருந்து பகிர்ந்த அளவைக் கழித்தும், மேல் எல்லையுடன் கூட்டியும் உண்மை இடைவெளிப் பிரிவுகளை (True class inteival) அமைக்கலாம்.

0—4		0·0—4·5	
5—9	என்ற உள்ளடக்கிய இடைவெளிகள்	4·5—9·5	என்ற உண்மை இடைவெளிப் பிரிவுகளாக அமைகின்றன.
10—14		9·5—14·5	
15—19		14·5—19·5	
20—24		19·5—24·5	

சில மாறிகள் தொடர்மாறிகளாக இருந்தபோதிலும் அளக்கும் கருவிகளின் தன்மைத்தாலும் நமது தேவையின் அடிப்படையாலும் முழு எண்களாகக் குறிக்கப்படுகின்றன. இவ் விடங்களில் அமைக்கப்பட்ட உள்ளடக்கிய இடைவெளிகளை உண்மை இடைவெளிகளாக மாற்றி அமைத்துக் கணக்கிடுவது நன்மை பயப்பதாகும். தொடர்மாறியான உயரம் 5′6″., 6′2″., 6′9″., என முழு எண்களாகக் குறிக்கப்படுவதால், இதற்காக அமைக்கப்படும் சிக்கலில்லாத உள்ளடக்கிய இடைவெளிகள் உண்மை இடைவெளிகளாக மாற்றி அமைக்கப்படவேண்டும்.

சம அளவுடைய மற்றும் சம அளவில்லாத (உள்ளடக்கிய, நுழைவு தவிர்க்கின்ற) இடைவெளிகளே நடைமுறையில் பெரும் பாலும் கையாளப்பட்டாலும் ஒருவர்தம் தேவைக்கேற்பக்

கட்டுப்பாடின்றிச் செயல்படலாம். குழந்தை இறப்பு விகிதம் (Infant mortality rate) அல்லது சிறுவர் இறப்பு விகிதம் (child mortality rate) படிக்க உள்ள இடத்தில் நமது தேவைக் கேற்பப் பிரிவு இடைவெளிகள் அமைத்துச் செய்திகள் தொகுக்கப்படலாம்.

பிரிவு இடைவெளி	இறந்தவர் எண்ணிக்கை
1 முதல் 7 நாள்கள்	...
8 முதல் 14 நாள்கள்	...
15 முதல் 21 நாள்கள்	...
22 முதல் 28 நாள்கள்	...
29 முதல் முதலாண்டு முடிய	...
இரண்டாவதாண்டு முதல் ஐந்தாம் ஆண்டு முடிய	...
மொத்தம்	...

பிரிவு இடைவெளியின் தன்மையும், எண்ணிக்கையும் நமது தேவையையும், நமக்குக் கொடுக்கப்பட்டுள்ள மாறிகளின் தன்மையையும் பொருத்து அமையும். இரு நிகழ்வெண் பரவல் களை ஒப்பிடும் போதும், நிகழ்வெண் பரவலிலிருந்து புள்ளியியல் அளவைகளை மதிப்பீடு செய்ய நேரிடும்போதும் முடிந்த அள விற்குச் சமமான இடைவெளிப் பிரிவுகள் கொண்ட நிகழ்வெண் பரவல்களை அமைத்தல் நலம். மாறிகளின் தன்மைக்கேற்பவும் குறிப்பிட்ட இடைவெளிகளில் அதிக நிகழ்வெண்கள் அமையும் தன்மைத்தாய் உள்ள இடங்களிலும் சமமில்லா இடைவெளிகளை அமைப்பதே நன்மை பயப்பதாகும். பொதுவாகப் பிரிவு இடை வெளிகளின் எண்ணிக்கை இருபதிற்குமேற் செல்லாமலும் பத்திற்குக் குறையாமலும் இருத்தல் நலம்.

பிரிவு இடைவெளிகளை அமைத்த பின், கொடுக்கப்பட்ட தொகுப்பிலுள்ள ஒவ்வோர் அளவும் அது சார்ந்துள்ள இடை

வெளியில் ஒரு சரிபார்க்கும் குறியால் (Tally mark—/) குறிக்கப்படும். நான்கு எண்களை ஒரு இடைவெளியில் அமைத்த பின், ஐந்தாவது எண்ணை, நான்கு எண்களின் சரிபார்க்கும் குறிகளின் மேல் குறுக்காக (卌) அமைக்கிறோம். அளவுகள் நிறைய உள்ள இடத்து, இறுதியில் நிகழ்வெண்களின் மதிப்பைக் கணக்கிட இக் குறியீட்டு முறை உதவியாய் உள்ளது. பட்டியலின் இறுதிக் கட்டத்தில் இச் சரிபார்க்கும் குறிகளின் கூட்டுத் தொகையைக் குறிப்பிடவேண்டும். ஆய்வினுக்கு எடுத்தாளப்படும் நிகழ்வெண் பரவல்களில் பிரிவு இடைவெளிகளும், அதன் நிகழ்வெண்களுமே தரப்படும். இச் சரிபார்க்கும் குறிகள் நிகழ்வெண் பரவல் அமைக்கும் இடத்தில்தான் குறிக்கப்படுகின்றன.

காட்டு: எண்பது பேர் உயரங்கள் கொடுக்கப்பட்டுள்ள பட்டியலுக்கு நிகழ்வெண் பரவல் அமைக்கும் முறையைப் பார்ப்போம்.

நிகழ்வெண் பரவல் அமைப்பு முறை

64·2	67·9	68·6	69·1	69·5	70·2	71·0	72·0
65·8	67·9	68·7	69·1	69·6	70·3	71·0	72·1
66·3	68·2	68·7	69·1	69·7	70·4	71·1	72·2
66·6	68·2	68·8	69·2	69·8	70·4	71·2	72·4
66·9	68·3	68·8	69·3	69·9	70·5	71·4	72·5
67·0	68·4	68·9	69·3	69·9	70·5	71·4	72·8
67·4	68·5	68·9	69·4	70·0	70·6	71·4	73·0
67·5	68·5	69·0	69·5	70·0	70·6	71·6	73·4
67·6	68·6	69·0	69·5	70·1	70·9	71·7	74·2
67·8	68·6	69·0	69·5	70·1	70·9	71·8	74·9

இவ் வளவுகளின் பெருமம் (maximum) 74·9 ஆகவும் சிறுமம் (minimum) 64·2 ஆகவும் அமைந்துள்ளன. இதன் வீச்சு (range) = 10·7· (74·9 − 64·2). ஒவ்வொரு இடைவெளியின் அளவும் ஓர்

பிரிவு இடைவெளி (x-உயரம்)	சரிபார்க்கும் குறி	நபர்கள் (நிகழ்வெண்)
64·0—64·9	1	1
65·0—65·9	1	1
66·0—66·9	111	3
67·0—67·9	~~1111~~ 11	7
68·0—68·9	~~1111~~ ~~1111~~ ~~1111~~	15
69·0—69·9	~~1111~~ ~~1111~~ ~~1111~~ 1111	19
70·0—70·9	~~1111~~ ~~1111~~ 1111	14
71·0—71·9	~~1111~~ ~~1111~~	10
72·0—72·9	~~1111~~ 1	6
73·0—73·9	11	2
74·0—74·9	11	2
	மொத்தம்	80

அங்குலமாகக் கொண்டால் பதினேர் இடைவெளிகளில் நம் அளவுகளின் மதிப்புகள் அடங்கும் என்பதைக் காணமுடியும். முதல் இடைவெளி சிறுமத்தை உட்கொள்வதாயும், இறுதி இடைவெளி பெருமத்தை உட்கொள்வதாயும் இருக்கவேண்டும். ஓர் அங்குல அளவுகொண்ட உள்ளடக்கிய பதினேர் இடைவெளிகளை அமைத்து ஒவ்வோர் அளவாகக் குறியிட்டால் நமக்குத் தேவையான நிகழ்வெண்பரவல் கிடைக்கும்.

ஒப்பீட்டு நிகழ்வெண் பரவல் (Relative frequency distribution)

சில இடங்களில், ஒவ்வொரு பிரிவு இடைவெளியிலும் அமையும் நிகழ்வெண்களைக் குறிப்பிடுவதைக் காட்டிலும், இவ் வெண்களுக்கும் மொத்த நிகழ்வெண்ணிற்கும் உள்ள விகிதத்தைக் குறிப்பிடுவது அவசியமாக உள்ளது. இவ் விகிதசமங்களைக் கொண்டு அமைக்கப்படும் நிகழ்வெண் பரவல் ஒப்பீட்டு நிகழ்

வெண் பரவல் (Relative frequency distribution) எனப்படும். மொத்த நிகழ்வெண்களை ஒன்று எனக் கொள்வதைத் தவிர்த்து நூறு எனக்கொண்டு நிகழ்வெண்களின் சதவீதங்களைக் குறிப்பிடுவதும் உண்டு. இத்தகைய பரவலுக்குச் சதவீத நிகழ் வெண் பரவல் (Percentage frequency distribution) என்று பெயர். இப் பரவல்கள் இவை பெறப்பட்ட எண்களிலிருந்து சார் பற்ற தன்மையை அடைகின்றன. கொடுக்கப்பட்ட நிகழ்வெண் பரவலுக்கில்லாத ஒரு பொதுத் தன்மை இந் நிகழ்வெண் பரவல் களுக்கிருப்பதை நம்மால் உணர முடிகிறது. மேலே கண்ட நிகழ் வெண் பரவலுக்கான ஒப்பீட்டு நிகழ்வெண் பரவல், சதவீத நிகழ்வெண் பரவல்கள் அடியிற்காணும் பட்டியலில் கணக்கிடப் பட்டுள்ளன.

பிரிவெண்	நிகழ் வெண்	விகிதசமம்	சதவீதம்
64·5	1	0·0125	1·25
65·5	1	0·0125	1·25
66·5	3	0·0375	3·75
67·5	7	0 0875	8·75
68·5	15	0·1875	18·75
69·5	19	0·2375	23·75
70·5	14	0·1750	17·50
71·5	10	0·1250	12·50
72·5	6	0·0750	7·50
73·5	2	0 0250	2·50
74·5	2	0 0250	2 50
	80	1·0000	100·00

வளர் நிகழ்வெண் பரவல் (Cumulative frequency distribution)

இன்னொரு முறையாகப் பெறப்படும் நிகழ்வெண் பரவல் வளர் நிகழ்வெண் பரவல் (Cumulative frequency distribution) எனப்படும். நிகழ்வெண் பரவலில் மாறியின் இரு மதிப்புகளுக்கிடையில் எவ்வளவு உறுப்புகள் அமைகின்றன என்று குறிப்பிடுகிறோம். ஒப்பீட்டு நிகழ்வெண் பரவலில் பிரிவு இடைவெளிகளில் உள்ள நிகழ்வெண்களுக்கும் மொத்த நிகழ்வெண்ணிற்கும் உள்ள விகிதசமத்தையோ அல்லது சதவீதத்தையோ குறிப்பிடுகிறோம்

இடைவெளிப் பிரிவு	நிகழ் வெண்	வளர் நிகழ்வெண் பரவல்			
		இடைவெளிப் பிரிவு மேல் எல்லை	கீழின வளர் நிகழ் வெண்	இடைவெளிப் பிரிவு கீழ் எல்லை	மேலின வளர் நிகழ் வெண்
64·0—64·9	1	64·9	1	64·0	80
65·0—65·9	1	65·9	2	65·0	79
66·0—66·9	3	66·9	5	66 0	78
67·0—67·9	7	67 9	12	67·0	75
68·0—68·9	15	68·9	27	68·0	68
69·0—69·9	19	69·9	46	69 0	53
70·0—70·9	14	70·9	60	70·0	34
71·0—71·9	10	71·9	70	71·0	20
72·0—72·9	6	72·9	76	72·0	10
73·0—73·9	2	73·9	78	73·0	4
74·0—74·9	2	74·9	80	74·0	2
	80				

வளர் நிகழ்வெண் பரவலில் ஒரு குறிப்பிட்ட மதிப்பிற்கு (பொதுவாக நிகழ்வெண் பரவலின் பிரிவு இடைவெளிகளின் கீழ் அல்லது மேல்எல்லை) மேல் அல்லது கீழ் உள்ள நிகழ்வெண்களின் கூட்டுத் தொகையைக் குறிப்பிடுகிறோம். இந் நிகழ்வெண் பரவல் கொடுக்கப்பட்ட நிகழ்வெண் பரவலின் மேலிருந்து அல்லது கீழிருந்து ஒவ்வொரு நிகழ்வெண்ணைச் சேர்த்து அமைக்கப் படுகிறது. மேலிருந்து கூட்டுவதால் கிடைக்கும் பரவல் கீழின வளர் நிகழ்வெண் பரவல் (less than cumulative frequency distribution) என்றும், கீழிருந்து கூட்டுவதால் கிடைக்கும் நிகழ்வெண் பரவல் மேலின வளர் நிகழ்வெண் பரவல் (greater than cumulative frequency distribution) என்றும் சொல்லப்படும். 80 பேர் உயரங்களுக்கான நிகழ்வெண் பரவல் முன்பக்கத்தில் உள்ள பட்டியலில் கீழினம் மற்றும் மேலின வளர் நிகழ்வெண் பரவல்களாக அமைக்கப்பட்டுள்ளது.

நிகழ்வெண் பரவல்களின் வரைபடங்கள் (Graphical representation of the frequency distribution)

கொடுக்கப்பட்ட அளவுகளின் பொது அமைப்பை நமக்கு எளிதாகத் தெரியப்படுத்த நிகழ்வெண் பரவல்களின் வரைபடங்கள் உபயோகப்படுகின்றன. ஒரு நிகழ்வெண் பரவலுக்கு அடியிற்காணும் வரைபடங்கள் அமைப்பது அதன் தன்மையைப் புரிந்துகொள்ள உதவியாக இருக்கும்.

1. நிகழ்வெண் செவ்வகப்படம் (Histogram).
2. நிகழ்வெண் பலகோணம் (Frequency Polygon).
3. நிகழ்வெண் வளைகோடு (Frequency curve).
4. வளர் நிகழ்வெண் வளைகோடு (cumulative Frequency curve).

நிகழ்வெண் செவ்வகப்படம் (Histogram) :

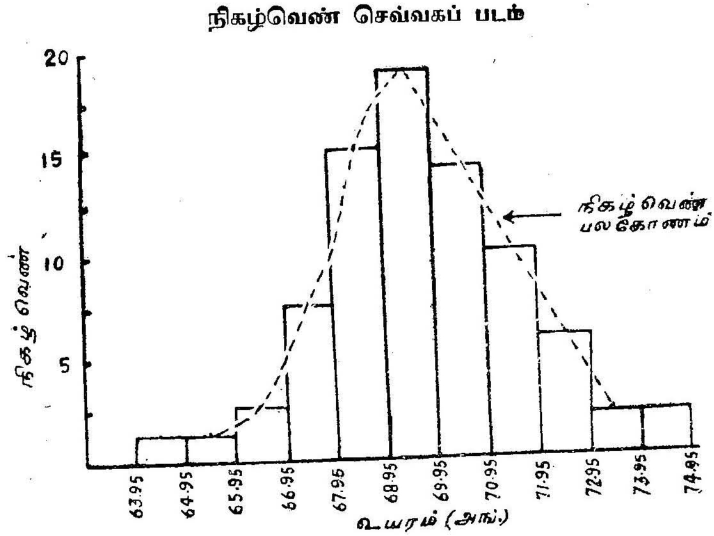
உண்மை இடைவெளிகளுக்கு மேல் அவ் விடைவெளிகளில் அமைந்த நிகழ்வெண்களின் மதிப்பிற்கு விகிதசமமான (Proportional) உயரங்களைக் கொண்ட செவ்வகங்களைத் தொடர்ச்சியாக அமைத்தால் நிகழ்வெண் செவ்வகப்படம் கிடைக்கும். x-அச்சில் (x-axis) பிரிவு இடைவெளிகளுக்கான சரியான அளவுத் திட்டத்தாலும், y-அச்சில் (y-axis) நிகழ்வெண்களுக்கான அளவுத் திட்டத்தாலும் இடைவெளியின்றி அமைக்கப்படும் செவ்வகங்கள் நிகழ்வெண் செவ்வகம் (Histogram) எனப்படும். ஒவ்வொரு செவ்வக

மும் அதன் பிரிவு இடைவெளியில் அமைந்த நிகழ்வெண்களுக்குச் சமமான பரப்பைக் கொண்டதாய் அமையவேண்டியது மிக முக்கியமானதாகும். சம அளவு இடைவெளிகளேக் கொண்ட பிரிவுகள் (class intervals of equal size) உள்ள நிகழ்வெண் பரவலுக்கான நிகழ்வெண் செவ்வகப்படம் (Histogram) பிரிவில் அமைந்த நிகழ்வெண்களின் மதிப்பிற்கேற்ற உயரங்களேக் கொண்டதாய் இருக்கும். சம அளவில்லாத பிரிவு இடைவெளிகளின் (cass intervals of unequal size) நிகழ்வெண் பரவலுக்கான நிகழ்வெண் செவ்வகப்படத்தின் செவ்வகங்கள் அப் பிரிவில் உள்ள நிகழ்வெண்களேப் பிரிவின் இடைவெளி அளவால் வகுத்து வரும் மதிப்புகளுக்குத் (Frequency/size of the class) தக்க உயரங்களேக் கொண்டதாயும் இருத்தல் வேண்டும். தனித்த மாறிகளுக்கான உள்ளடக்கிய இடைவெளிப் பிரிவுகளே (Inclusive class intervals *0—4, 5—9, 10—14*...etc.,), உண்மை இடைவெளிப் பிரிவுகளாக மாற்றி (True class intervals, *0—4·5, 4·5—9·5, 9·5—14·5*...etc.,) செவ்வகங்கள் வரைய வேண்டும். செவ்வகங்கள் தொடர்ச்சியாக அமையவேண்டி இருப்பதால் உள்ளடக்கிய இடைவெளிகளே உண்மை இடைவெளிகளாக மாற்றியமைக்க வேண்டிய நிலே ஏற்படுகிறது.

காட்டு :

80 பேர் உயரங்களுக்காக அமைக்கப்பட்ட நிகழ்வெண் பரவலின் இடைவெளிப் பிரிவுகளுக்கு உண்மை இடைவெளிகளே அமைத்து நிகழ்வெண் செவ்வகம் வரைவோம்.

உண்மை இடைவெளிப் பிரிவு	63·95 64·95	64·95 65·95	65·95 66·95	66·95 67 95	6 7·95 68·95	68·95 69·95	69·95 70·95	70·95 71·95	7 1·95 72·) 3	72·95 73 95	73·95 74·95
நிகழ்வெண்	1	1	3	7	15	1[illegible]	14	1 0	6	2	2



படம் 2·1

x-அச்சில் பிரிவு இடைவெளிகளையும் (உயரம்), y- அச்சில் நிகழ்வெண்களையும் கொண்டு நிகழ்வெண் செவ்வகப்படம் வரைந்துள்ளோம்.

காட்டு :

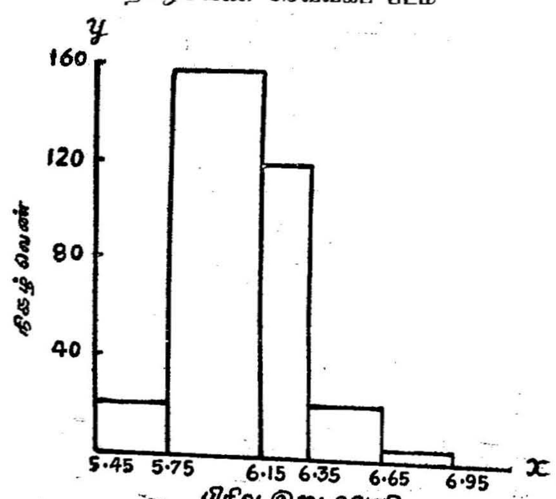
பிரிவு இடைவெளி	நிகழ்வெண்
5·5—5·7	58
5·8—6·1	632
6·2—6·3	241
6·4—6·6	64
6·7—6·9	5
	1000

கொடுக்கப்பட்ட நிகழ்வெண் பரவலின் இடைவெளிப் பிரிவுகள் சம அளவில்லாத உள்ளடக்கிய இடைவெளிப் பிரிவு களாக அமைந்துள்ளன. நிகழ்வெண் செவ்வகத்திற்குத் தேவை யான உண்மை இடைவெளிப் பிரிவுகளை முதற்படியாக அமைக்

கிறோம். உண்மை இடைவெளிப் பிரிவுகளுக்குமேல், பிரிவு நிகழ்வெண்களின் மதிப்பிற்கு ஏற்றுற்போல் செவ்வகங்களின் உயரங்களை அமைக்காமல், இந் நிகழ்வெண்களைப் பிரிவு இடைவெளி அளவுகளால் வகுத்து வரும் மதிப்பிற்கு ஏற்றுற்போல் உயரங்களை அமைக்கவேண்டும். இவ்வாறு அமைக்கப்படும் நிகழ்வெண் செவ்வகப்படமே, பிரிவு இடைவெளிகளின் மேல் அமையும் செவ்வகங்கள், அவ் விடை வெளியில் அமையும் நிகழ்வெண்களின் பரப்பிற்குச் சமமாக அமையும்.

உண்மை இடைவெளி (True class interval)	நிகழ் வெண்	நிகழ்வெண் ÷ பிரிவு அளவு	
5·45—5·75	58	58 ÷ 3	19.33
5·75—6·15	632	632 ÷ 4	158·00
6·15—6·35	241	241 ÷ 2	120·50
6·35—6·65	64	64 ÷ 3	21·33
6·65—6·95	5	5 ÷ 3	1·67

நிகழ்வெண் செவ்வகப் படம்



படம் 2·2

இடைவெளிப் பிரிவுகள் சமமாக இல்லாத இடங்களிலும், விரைவாக மாறும் தன்மைத்தான நிகழ்வெண்களைக் கொண்ட மாறிக்கான நிகழ்வெண் பரவல்களின் தன்மையை நன்குணரவும் நிகழ்வெண் செவ்வகப்படம் உதவிசெய்கிறது.

நிகழ்வெண் பலகோணம் (Frequency polygon):

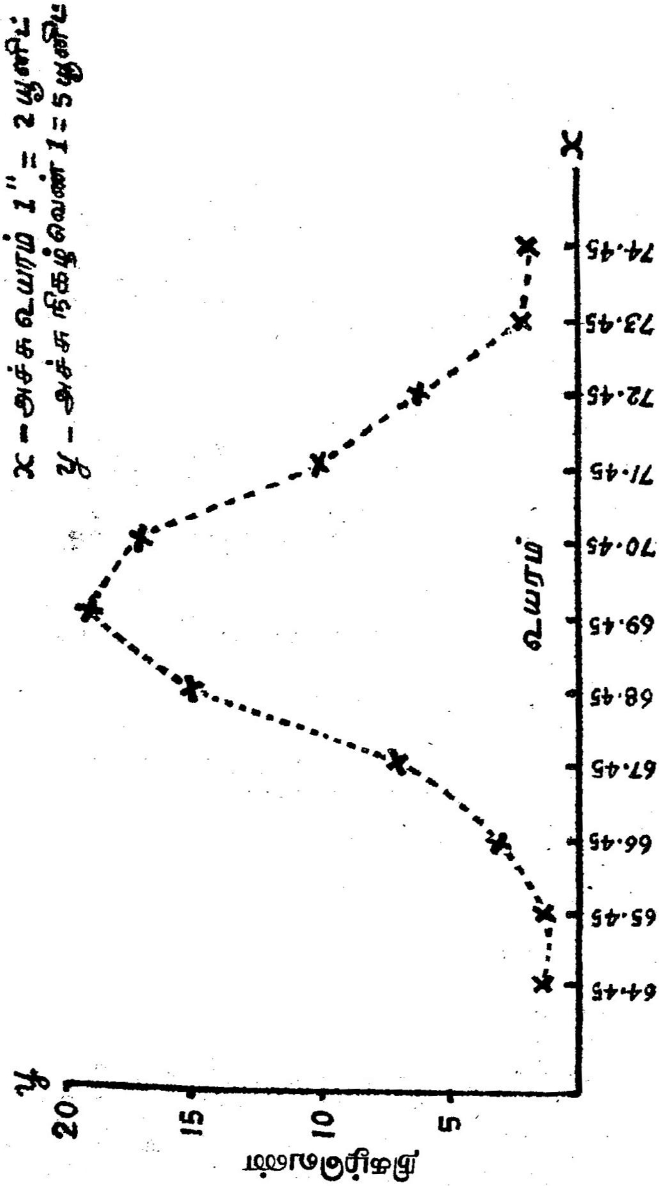
இரண்டு நிகழ்வெண் பரவல்களை ஒப்பிடும்போது இரு நிகழ்வெண் செவ்வகப்படங்களை ஒரே படத்தில் வரையும் நிலையாய் அமைகிறது. இத் தன்மைத்தான இடங்களில் நிகழ்வெண் செவ்வகப்படங்களுக்குப் (Histogram) பதிலாக நிகழ்வெண் பலகோணங்களை வரைவதால், இரண்டு அல்லது மூன்று நிகழ்வெண் பரவலுக்கான நிகழ்வெண் பலகோணங்கள் ஒரே படத்தில் வரையப்பட்டு ஒப்பிடும் முறை சிறந்ததாக்கப்படும். இடைவெளிப் பிரிவுகளின் பிரிவெண்களை (mid x) x-அச்சிலும், பிரிவு நிகழ்வெண்களை (class frequency) y-அச்சிலும் குறித்து நிகழ்வெண் பல கோணம் வரையப்படுகிறது. பிரிவெண்களுக்கு நேராகக் குறிக்கப்பட்ட புள்ளிகளை நேர்கோடுகளால் சேர்த்து நிகழ்வெண் பலகோணம் அமைக்கப்படுகிறது.

நிகழ்வெண் செவ்வகப்படத்தின் செவ்வகங்களின் மையப் புள்ளிகளை நேர்கோடுகளால் சேர்ப்பதாலும் நிகழ்வெண் பல கோணம் கிடைக்கும். 80 பேர் உயரங்களுக்காக அமைக்கப்பட்ட நிகழ்வெண் பரவலின் செவ்வகப்படத்தில் (- - - -) குறியிட்ட அமைப்பே நிகழ்வெண் செவ்வகப்படத்திலிருந்து செவ்வகங்களின் மையப் புள்ளியைச் சேர்ப்பதால் கிடைக்கும் நிகழ்வெண் பல கோணமாகும் (படம் 2·1). அந் நிகழ்வெண் பரவலுக்கான நிகழ்வெண் பலகோணம் தனியாக வரையப்பட்டுள்ளதை அடுத்த பக்கத்தில் காணலாம்.

நிகழ்வெண் வளைகோடு (Frequency Curve)

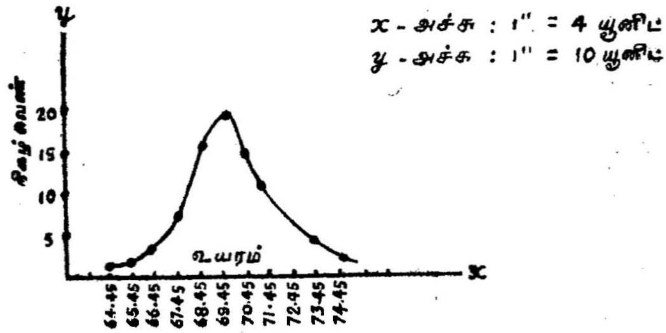
நிகழ்வெண்களின் எண்ணிக்கையை அதிகப்படுத்தியும் பிரிவு இடைவெளிகளைச் சிறியதாக்கியும் செய்யும்போது நிகழ்வெண் செவ்வகமும் நிகழ்வெண் பலகோணமும் ஒரு சீரான வளை கோடாக அமையும். இவ் இறுதித் தன்மைத்தான வளைகோட்டின் அமைப்பே நிகழ்வெண் வளைகோடெனப்படும். இதன் அமைப்பிற்கான ஓர் உருவைக் (model) காண்பதும், அதன் குணங்களைக் கொண்டு நமது மாதிரி எடுக்கப்பட்ட முழுமைத் தொகுதியின் தன்மைகளைக் கூறுவதுமே புள்ளியியலின் முக்கிய நோக்கமாகும்.

நிகழ்வெண் பலகோணம்



படம் 2.3

நிகழ்வெண் வளைகோடு

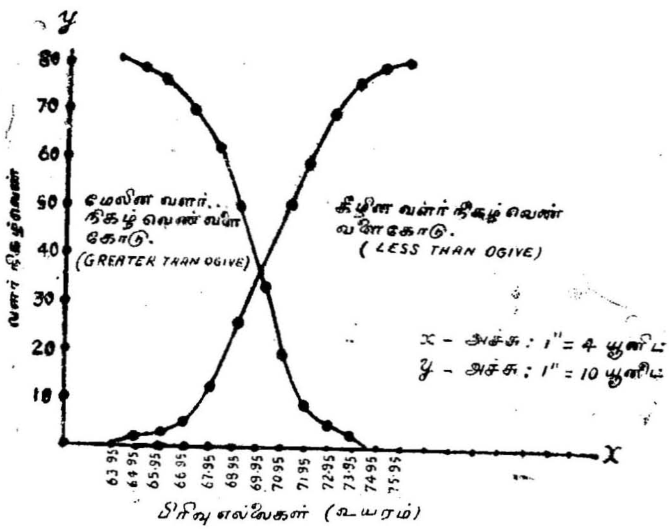


படம் 2.4

வளர் நிகழ்வெண் வளைகோடு (Cumulative Frequency curve) :

நிகழ்வெண் கீழினச் சேர்க்கைக்கும் மேலினச் சேர்க்கைக்கும் வரையப்படும் நிகழ்வெண் வளைகோடுகளே கீழின வளர் நிகழ்

வளர் நிகழ்வெண் வளைகோடுகள்



படம் 2·5

வெண் வளைகோடு (less than cumulative frequency curve) எனவும் மேலின வளர் நிகழ்வெண் வளைகோடு (greater than cumulative frequency curve) எனவும் சொல்லப்படும். உண்மைப்

பிரிவு எல்லைகளை x—அச்சில் அமைத்து இவற்றிற்கு நேராகக் கீழினம் அல்லது மேலின வளர் நிகழ்வெண்களை y—அச்சில் குறித்து முறையான புள்ளிகளைச் சேர்த்து வளர் நிகழ்வெண் வளைகோடுகள் வரையப்படுகின்றன. கீழின வளைகோடுகளுக்கான புள்ளிகள் பிரிவு எல்லைகளின் மேல் எல்லைக்கு நேராகவும், மேலின வளைகோட்டிற்கான புள்ளிகள் பிரிவு எல்லைகளின் கீழ் எல்லைகட்கு நேராகவும் குறிக்கப்படவேண்டியது உணரத்தக்கது. இவையே கீழினம் மற்றும் மேலின ''ஒகைவ்'' (less than or greater than ogive) என ஆங்கிலத்தில் கூறப்படும்.

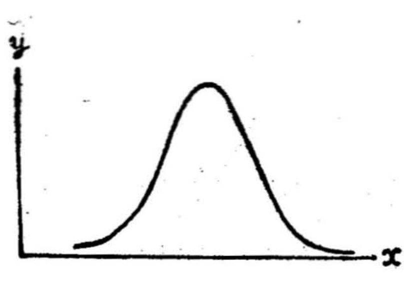
நிகழ்வெண் வளைகோடுகளின் சில முக்கியமான வகைகள் :

நிகழ்வெண் பரவல்களை (பெரும்பாலும்) நான்கு வகையாகப் பிரிக்கலாம்.

1. சமச்சீர் பரவல் (Symmetrical Distribution)
2. கோட்டமுடைப் பரவல் (Skewed Distribution)
3. J-வடிவப் பரவல் (J-shaped Distribution)
4. U-வடிவப் பரவல் (U-shaped Distribution)

வேறுவிதமாக அமையக் கூடிய பரவல்கள் கிடைக்கப் பெற்றாலும், கணக்கியல் உதவி கொண்டு, இந் நான்கு முக்கிய மான வடிவங்களில் ஏதாவதொன்றைச் சார்ந்ததாகக் கொள்ள முடியும்.

சமச்சீர் வளைகோடு (Symmetrical curve):

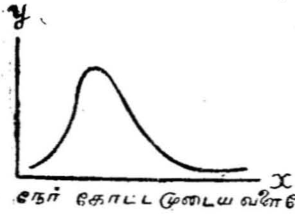


படம் 2·6

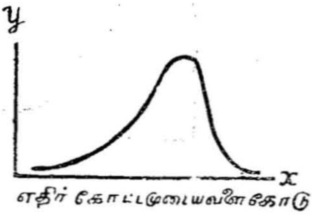
இத் தன்மைத்தான பரவலில், பிரிவு நிகழ்வெண்கள் மையத்திலிருந்து இருபுறங்களிலும் சீராகக் குறைந்து பூச்சியத்தை அடையும். உயிர்ப் புள்ளியியலில் இத் தன்மைத்தான் பரவல்கள் அமையக்கிடைப்பது மிகவும் குறிப்பிடத்தக்க ஒன்றாகும். இதன்

வரைபடம் சரியாக மடிக்கப்பட்டால், மையத்திலிருந்து இருபுறமும் சரியாக இணைவதைக் காணலாம். இதுவே மணி (∩) போன்ற (bell shaped) உருவுடைய வளைகோடெனவும் சொல்லப்படுகிறது.

கோட்டமுடைப் (மிதமான சமச்சீரற்ற) பரவல்கள் (Skewed curves) :



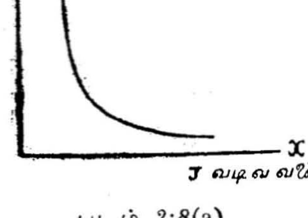
படம் 2·7(a)

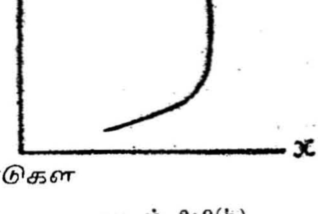


படம் 2·7(b)

சமச்சீராக அமையாத பரவல்கள் கோட்டமுடைப் பரவல்கள் எனச் சொல்லப்படும். நடைமுறையில் கிடைக்கக்கூடிய பெரும்பாலான பரவல்கள் இம்முறையில் அமையப்பெறும். இம்மாதிரிப் பரவல்களில் பிரிவு நிகழ்வெண்கள் மையக் கோட்டிலிருந்து இருபுறமும் சமமாகப் பரவாமல், ஒரு புறத்தைச் சார்ந்து இருக்கும். இதற்கான வளைகோட்டின் உச்சத்திலிருந்து நீண்ட வால்புறம் வலப்புறம் அமையும்போது நேர்கோட்டமுடை (positively skewed) வளைகோடெனவும், உச்சத்திலிருந்து நீண்ட வால்புறம் இடப் புறத்தில் அமையும்போது எதிர்கோட்டமுடை (negatively skewed) வளைகோடெனவும் சொல்லப்படும்.

J-வடிவப் (J-shaped) பரவல்கள் :



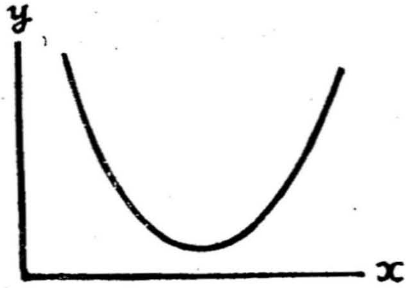


J வடிவ வளைகோடுகள்

படம் 2·8(a) படம் 2·8(b)

இம்மாதிரியான பரவல்களில் நிகழ்வெண்கள் மாறியின் இறுதி மதிப்பில் மிகுந்து அமைகின்றன. இவற்றின் இறுதி வடிவம் 'J' போன்று அமைவதால், இவை 'J' வடிவ வளை கோடுகள் எனவும் சொல்லப்படுகின்றன.

'U'-வடிவ வளைகோடுகள் (U-shaped curves) :



U வடிவ வளைகோடுகள்

படம் 2·9

மையத்தில் குறைந்த நிகழ்வெண்களும், மாறியின் இறுதி மதிப்பில் (இருபுறமும்) அதிகமான நிகழ்வெண்களும் அமையப் பெற்ற பரவல்கள் U-வடிவப் பரவல் எனப்படும். இம் மாதிரிப் பரவல்கள் மிகவும் அரிதாகப் பெறப்பட்டாலும், மேற்கண்ட வற்றினின்றும் மாறுபட்டு ஒரு தனித் தன்மையுடன் அமைவதால் முக்கியத்துவம் பெறுகிறது.

பயிற்சி :

1. ஒரு கூடையிலிருந்து ராண்டமாக எடுக்கப்பட்ட ஆப்பிள் களின் எடை கிராமில் தரப்பட்டுள்ளது. ஆப்பிள்களின் எடையை ஒரு மாறியாகக் கொண்டு, ஒரு நிகழ்வெண் பரவலை அமை.

106 107 76 82 109 107 115 93 187 95 123 125 186

92 86 70 126 68 130 129 139 119 115 128 100 78

84 99 113 204 111 141 136 123 90 115 98 110

90 107 81 131 75 84 104 110 80 118 82 111

2. அடியிற்காணும் நிகழ்வெண் பரவலுக்கான வளர் நிகழ் வெண் பரவல்களை அமைத்து, வளர் நிகழ்வெண் வளைகோடுகளை அமை.

பிரிவு இடைவெளி	0—4	4—8	8—12	12—16	16—20
நிகழ்வெண்	4	6	10	8	4

3. அடியிற் காணும் நிகழ்வெண் பரவலுக்கான உண்மை இடைவெளிப் பிரிவுகளை (True class Intervals) அமைத்து (1) நிகழ்வெண் செவ்வகப்படம் (2) நிகழ்வெண் பலகோணம் ஆகியவற்றை வரைக.

பிரிவு இடைவெளி	நிகழ்வெண்
60—64	2
65—69	15
70—74	67
75—79	162
80—84	331
85—89	567
90—94	827
95—99	1094
100—104	1240
105—109	1363
110—114	1086
115—119	755
120—124	598
125—129	345
130—134	190
135—139	59
140—144	31
145—149	8
மொத்தம்	8740

4. கொடுக்கப்பட்டுள்ள அளவுகளுக்கான நிகழ்வெண் பரவலை அமைத்து நிகழ்வெண் வளைகோட்டை (Frequency curve) வரை.

45·7	41·9	40·2	38·9	40·0	39·5	45·5	47·4
47·1	42·7	51·0	48·0	47·6	50·0	52·6	50·1
54·1	57·0	54·2	58·1	61·7	60·2	58·7	61·0
63·8	63·6	62·2	57·2	56·9	54·2	52·4	53·6
46·0	49·8	45·4	40·1	47·2	45·3	46·2	40·4
49·8	46·6	59·7	55·2	59·8	58·8	51·5	48·8
42·4	49·0	44·4	81·1				

5. அடியிற்காணும் வளர் நிகழ்வெண் பரவலுக்கான (cumulative frequency distribution) நிகழ்வெண் பரவலை அமைத்து நிகழ்வெண் செவ்வகம் (Histogram) வரை.

பிரிவு இடைவெளியின் மேல் எல்லை	வளர் நிகழ்வெண்
5	9
15	22
25	108
30	208
45	347
55	467
60	513
65	525

6. கொடுக்கப்பட்டுள்ள வளர் நிகழ்வெண் பரவலுக்கான நிகழ்வெண் பரவலையும், நிகழ்வெண் வளைகோட்டையும் அமை.

மாறியின் கீழ் எல்லை	0	10	20	30	40	50	60	70	80
நிகழ்வெண்	290	289	282	254	176	92	37	9	2

7. மேலின மற்றும் கீழின வளர் நிகழ்வெண் பரவல், சதவீத வளர் நிகழ்வெண் பரவல்கள் கொடுக்கப்பட்டுள்ளன. இதற்கான ஓகைவுகள் (Ogives) வரை. நிகழ்வெண் பரவலையும், சதவீத நிகழ்வெண் பரவலையும் அமை.

இடைவெளிப் பிரிவு	கீழின வளர் நிகழ்வெண்	மேலின வளர் நிகழ்வெண்	சதவீத கீழின வளர் நிகழ்வெண்	சதவீத மேலின வளர் நிகழ்வெண்
72·0—73·9	7	225	3·1	100·0
74.0—75·9	38	218	16·9	96·9
76·0—77·9	80	187	35·6	83·1
78·0—79·9	134	145	59·6	64·4
80·0—81·9	167	91	74·2	40·4
82·0—83·9	191	58	84·9	25·8
84·0—85·9	213	34	94·7	15·1
86·0—87·9	221	12	98·2	5 ·3
88·0—89·9	225	4	100·0	1·8

8. 1000 மாணவர்களின் உயரம் செ.மீ. கொடுக்கப்பட்டுள்ளது. 'இதற்கான உண்மை இடைவெளிப் பிரிவுகளைக் குறித்து நிகழ்வெண் செவ்வகம் வரை.

உயரம் செ.மீ.	மாணவர்கள் எண்ணிக்கை
155—157	4
158—160	8
161—163	26
164—166	53
167—169	89
170—172	146
173—175	188
176—178	181
179—181	125
182—184	92
185—187	60
188—190	22
191—193	4
194—196	1
197—199	

9. கொடுக்கப்பட்ட நிகழ்வெண் பரவலுக்கான நிகழ்வெண் செவ்வகம், நிகழ்வெண் பலகோணம் வரை.

x	1	2	3	4	5	6	7	8	9	10	11	12
f	1	7	11	16	8	4	5	2	1	0	0	1

3. நிகழ்வெண் பரவல் வருணனைகள்
(Description of a Frequency Distribution)

சராசரிகள் (Averages) :

தொகுத்த செய்திகளை நிகழ்வெண் பரவலாகச் சுருக்கி அமைத்தல், அவற்றைப் புரிந்து கொள்வதற்கான முதற்படியாகும். செய்திகளின் ஆய்வுக் கருத்துகள் நிறையிட்ட சொற்களால் சுருக்கமாக விளக்கப்பட்டாலன்றி அவற்றைச் செயல்படுத்தத் தகுதியாக்கிக் கொள்ளமுடியாது. எனவே செய்திகளின் முக்கியமான சிறப்பியல்புகளைக் குறிக்கும் அளவைகளைக் (statistic) காணும் தேவை ஏற்படுகிறது. இத்தகைய அளவைகள் வருணனை அளவைகள் எனப்படும். நிகழ்வெண் பரவலின் சிறப்புப் பண்புகளை வருணிப்பதே (characterise) இவ்வளவைகளின் முக்கிய நோக்கமாகும். கொடுக்கப்பட்ட செய்திகளின் பொது அளவைப் புரிந்து கொள்ளும்படியான எளிதான மற்றும் சுருக்கமான எண்களால் செய்திகளை விவரிப்பதே இவற்றின் முக்கிய நோக்கமாகும். ஒரே இனமான அளவுத் திரள்களின் முழுமைக்கும் ஒரு பிரதிநிதியாக விளங்குவதே இதன் அடிப்படைத் தத்துவமாகும். பொதுவாக இவ்வளவைகள் நிகழ்வெண் பரவலின் மத்திய பாகத்தில் காணப்படுவதால், இவ்வளவைகள் மைய நிலைப் போக்கைக் குறிக்கும் அளவைகள் (Measures of central tendency) எனவும் சொல்லப்படும். மைய நிலைப்போக்கு, மத்தியில், பிரதிநிதித்துவம் என்ற சொற்கள் பல்வேறுவிதமாக விளக்கப்படும் தன்மைத்தாய் அமைந்துள்ளன. ஈடாக மைய நிலைப்போக்கைக் குறிக்கவும் பல்வேறு விதமான சராசரி அளவைகள் (Averages) நமக்குக் கிட்டி உள்ளன.

சராசரி எனும் சொல் குறிக்கும் பண்புகளின் திரள் அனைத்தும் சமமான பொருள் தரா என்பது வெளிப்படை. நடைமுறையில் சராசரி (Average) எனும் சொல் கூட்டுச் சராசரியை

(Arithmetic mean) மட்டும் குறிப்பதாகக் கொள்ளப்படுகிறது புள்ளியியலில் சராசரி எனும் சொல் சராசரி அளவைகள் அனைத்தையும் குறிக்கும் ஒரு பொதுப்பெயராகும். இவ் வேறுபாட்டினை உணர்வது மிக இன்றியமையாததாகும். ஒரு தனிப்பட்ட சராசரியைக் குறிப்பிட்டுச் சொன்னாலன்றி, சராசரி எனும் சொல் ஒரு பொதுப் பெயராகவே கையாளப்படும். உயிர்ப் புள்ளியியலில் பெரும்பாலும் கையாளப்படும் மூன்று சராசரிகள்

1. கூட்டுச் சராசரி (Arithmetic mean)
2. இடை நிலை (Median)
3. முகடு (Mode)

என்பனவாகும். இவற்றிலும் கூட்டுச் சராசரியே மிகுதியும் பயன்படுத்தப்படுகிறது என்பது குறிப்பிடத்தக்கது. ஆய்விற்குரிய செய்திகளின் தன்மைக்கேற்பவும் சிக்கலை அணுகும் முறைக்கேற்பவும் நம் அனுபவத்தைக்கொண்டு சரியான சராசரியைத் தேர்ந்தெடுத்துக் கொள்ளவேண்டும்.

கூட்டுச் சராசரி (Arithmetic mean):

மொத்த அளவுகளின் கூட்டுத் தொகையை கூட்டுத் தொகையை ஏற்படுத்திய உறுப்புகளின் எண்ணிக்கையால் வகுக்கக் கிடைக்கும் மதிப்பே கூட்டுச் சராசரி எனப்படும். மாதிரியின் அளவுகளை $x_1\ x_2 \ldots\ldots x_n$ என்ற n உறுப்புகளாகக் கொண்டால் மொத்தக் கூட்டுத் தொகைக்கும் உறுப்புகளின் எண்ணிக்கைக்கும் உள்ள ஈவை $\bar{x}$ (x-பார்) எனக் குறிப்பிடுகிறோம்

$$\bar{x} = \frac{x_1 + x_2 + \ldots\ldots + x_n}{n} = \frac{\sum_{i=1}^{n} xi}{n}$$

Σ (Sigma) என்பது கிரேக்க எழுத்தாகும். மாறிகளின் கூட்டுத் தொகையைக் குறிக்க இக் குறியீட்டினை உபயோகப்படுத்துகிறோம். ஒன்று முதல் n வரை ($i = 1$ to n) எனக் குறிக்கப்பட்டுள்ளது, மாறியின் அளவுகள் ஒன்றிலிருந்து 'n' வரை (n-உறுப்புகளின் அளவுகள்) கூட்டப்படவேண்டும் என்பதைக் குறிக்கும். காட்டாக,

$\sum_{i:1}^{3} xi = x_1 + x_2 + x_3$ என்பதையும்

$$\sum_{i:1}^{5} xi = x_1 + x_2 + x_3 + x_4 + x_5$$ என்பதையும் குறிக்கும்.

Σ-எனும் குறியீடு கணக்கியல் முறையை மிகவும் எளிதாக்குகிறது.

$$\sum_{1}^{n} c\, xi = c\sum_{1}^{n} xi = c[x_1 + x_2 + \ldots\ldots + x_n]$$ என்பதையும்

$$\sum_{1}^{n} c = c + c + \ldots\ldots + c - n \text{ முறைகள்} = n.c$$ என்பதையும்

$$\sum_{i:1}^{n} xi^2 = x_1^2 + x_2^2 + \ldots\ldots + x^2_n$$ என்பதையும் குறிக்கும்.

காட்டு :

11 நபர்களின் உயரம் அங்குலத்திற்குச் சரியாக அளக்கப் பட்டு 63, 63, 64, 65, 66, 66, 68, 68, 68, 69, 69—என மதிப் பிடப்பட்டுள்ளன. இதன் கூட்டுச் சராசரியைக் கணக்கிடும் முறையைப் பார்ப்போம்.

இவ் வளவுகளை $x_1\ x_2 \ldots\ldots x_{11}$ எனக் கொள்வோம். உறுப்பு களின் மொத்த எண்ணிக்கை $(n) = 11$ என அமைகிறது.

$$\text{உயரங்களுக்கான கூட்டுச் சராசரி} = \frac{x_1 + x_2 + \ldots + x_{11}}{n}$$

$$= \frac{\sum xi}{n} = \frac{\sum_{1}^{11} xi}{11}$$

$$= \frac{63 + \ldots + 69}{11} = \frac{729}{11}$$

$$\overline{x} = 66{\cdot}3 \text{ அங்.}$$

சுலப முறை :

எண்ணிக்கையில் அதிகமாகவும் பெரியதுமான அளவுகள் உள்ள இடத்துக் கூட்டுச் சராசரியைக் கணக்கிடும்போது அடியிற் காணும் சுலபமுறையை உபயோகப்படுத்துவது கணக்கிடும் முறையில் தவறுகளையும் நேரத்தையும் குறைப்பதற்கு ஏதுவாய் உள்ளது. இம்முறையில் கொடுக்கப்பட்ட அளவுகளின் தன்மைக் கேற்ப ஒரு தோராயமான எண்ணைக் (மதிப்பை) கூட்டுச் சராசரி யாகக் கொண்டு, உண்மையான கூட்டுச் சராசரி கணக்கிடப் படுகிறது.

$x_1 x_2 \ldots x_n$ என்பன கொடுக்கப்பட்ட அளவுகளாக இருக் கட்டும். ஒவ்வொரு அளவிலிருந்தும் 'A'—என்ற தோராய மதிப்பைக் கழித்துவரும் அளவுகளின் மதிப்பை $d_1\ d_2 \ldots dn$ எனக் கொள்வோம். பெறப்பட்ட அளவுகளுக்கான $(d_1\ d_2 \ldots dn)$

கூட்டுச் சராசரியைக் ($\bar{d}$) கணக்கிட்டு, கொடுக்கப்பட்ட அளவுகளுக்கான ($x_1 x_2 \; x_n$) கூட்டுச் சராசரி ($\bar{x}$) மதிப்பிடப்படுகிறது.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{1:1}^{n} x i}{n}$$

$$di = xi - A.$$

$$\Sigma di = \Sigma (xi - A).$$

$$\Sigma di = \Sigma xi - \Sigma A$$

$$\frac{\Sigma di}{n} = \frac{\Sigma xi}{n} - \frac{\Sigma A}{n}$$

$$\bar{d} = \bar{x} - A.$$

$$\bar{x} = A + \bar{d}$$

11. நபர்களின் உயரங்களுக்கான அளவுகளுக்கு 66ஐத் தோராய மதிப்பாகக் கொண்டு கூட்டுச்சராசரியைக் கணக்கிடுவோம்.

உயரம் xi	$(x_i - A) = (x_i - 66)$ di +	−
63		3
63		3
64		2
65		1
66		0
66		0
68	2	
68	2	
68	2	
69	3	
69	3	
	12	9
	+ 3	

$$\Sigma di = +3$$

$$\bar{d} = \frac{3}{11} = 0{\cdot}27$$

$$A = 66$$

$$\bar{x} = A + \bar{d}$$

$$= 66 + 0{\cdot}27$$

$\bar{x} = 66{\cdot}27$ அங்

கூட்டுச்சராசரி : தனித்த மாறி (Discrete Variable) யின் பரவலுக்கானது

தனித்த மாறிகளைப் பகுக்கும்போது, அவற்றின் மதிப்பிற்குத் தக்கவாறு, ஒவ்வொரு மதிப்பிற்கும் உள்ள நிகழ்

வெண்கள் அதற்கெதிராகக் குறிக்கப்படும். $x_1, x_2 \dots x_n$ என்ற தனித்த மாறியின் (discrete variable) அளவுகள் $f_1\, f_2 \dots fn$ முறை நிகழும்போது

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots\dots + x_n fn}{f_1 + f_2 + \dots\dots + fn} = \frac{\Sigma\, xi\, fi}{\Sigma\, fi}$$

என அமையும்.

சுலப முறையில் கணக்கிடும்போது, ஒவ்வொரு மதிப்பிலிருந்தும் தோராய மதிப்பைக் (**A**) கழித்து வரும் விலக்கங்களை (*di*) அதற்கொத்த நிகழ்வெண்களால் (*fi*) பெருக்கிவரும் கூட்டுத் தொகையை, நிகழ் வெண்களின் கூட்டுத் தொகையால் வகுக்க வேண்டும்.

கூட்டுச் சராசரி : நேர்முறை

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots\dots + xn\, fn}{f_1 + f_2 + \dots\dots + fn} = \frac{\Sigma\, xifi}{\Sigma fi} = \frac{\Sigma\, xifi}{n}$$

கூட்டுச்சராசரி : சுலபமுறை

$$\bar{d} = \frac{d_1 f_1 + d_2 f_2 + \dots\dots + dn\, fn}{f_1 + f_2 + \dots\dots + fn} = \frac{\Sigma\, difi}{\Sigma fi} = \frac{\Sigma\, difi}{n}$$

$$di = xi - \text{A}$$

$$xi = \text{A} + di$$

$$\bar{x} = \text{A} + \bar{d}$$

நேர் முறை			சுலப முறை					
xi	fi	$xifi$	xi	fi	$xi - \text{A}$ di		$di\, fi$	
					+	−	+	−
x_1	f_1	$x_1 f_1$	x_1	f_1				
x_2	f_2	$x_2 f_2$	x_2	f_2				
.	.	.	.	.				
.	.	.	.	.				
.	.	.	.	.				
.	.	.	.	.				
.	.	.	.	.				
x_n	f_n	$x_n f_n$	x_n	f_n				
	$\Sigma fi = n$	$\Sigma xifi$		$\Sigma fi : n$				
							$\Sigma\, difi$	

இவ் வளவுகளைக் கணக்கிட உதவும் பட்டியல் 35 ஆம் பக்கத்தில் கொடுக்கப்பட்டுள்ளது.

காட்டு : 56 மரத் தூறுகளின் விட்டங்கள் (அடிகளில்) கொடுக்கப்பட்டுள்ளன. இதற்கான கூட்டுச் சராசரியைக் (நேர்முறை, சுலபமுறை) கணக்கிடுவோம்.

நேர்முறை			சுலபமுறை			
தூறுகளின் விட்டம் x_i	மரங்கள் f_i	$x_i \times f_i$	$x_i - 4$		$d_i \times f_i$	
			+	−	+	−
1	1	1		3		3
2	7	14		2		14
3	11	33		1		11
4	16	64		0		
5	8	40	1		8	
6	4	24	2		8	
7	5	35	3		15	
8	2	16	4		8	
9	1	9	5		5	
10	0	0	6		0	
11	0	0	7		0	
12	1	12	8		8	
	56	248			52	28
					+ 24	

கூட்டுச் சராசரி : நேர்முறை

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_n f_n}{f_1 + f_2 + \dots f_n} = \frac{\Sigma x_i f_i}{n}$$

$$= \frac{248}{56} = 4.4285$$

கூட்டுச் சராசரி : சுலபமுறை

$$\bar{d} = \frac{d_1 f_1 \ d_2 f_2 + \ldots\ldots + dn\,fn}{f_1 + f_2 + \ldots\ldots + fn} = \frac{\Sigma\, di\, fi}{\Sigma\, fi} = \frac{\Sigma\, di\, fi}{n}$$

$$= \frac{24}{56} = 0{\cdot}4285$$

$$\bar{x} = A + \bar{d} \quad [A = 4]$$

$$= 4 + 0{\cdot}4285$$

$$= 4{\cdot}4285$$

கூட்டுச் சராசரி : நிகழ்வெண் பரவலுக்கானது

கொடுக்கப்பட்ட அளவுகளை இடைவெளிப் பிரிவுகளில் சேர்த்து விட்டால், அவற்றின் மதிப்புகளை உறுதியாகக் கணக்கிட முடியாத நிலையாகி விடுகிறது. இவ்விடங்களில் ஒரு இடைவெளிப் பிரிவில் அமைந்த நிகழ்வெண்கள் அனைத்தும் இடைவெளிப் பிரிவின் மையப் புள்ளியில் (பிரிவெண்ணில்-mid point) குவிவதாகக் கொள்ளப்படும். அதாவது இடைவெளியில் அமைந்த அளவுகள் அனைத்தும் இடைவெளிப் பிரிவெண்ணிற்குச் சமமானதாகக் கொள்ளப்படும். பிரிவு இடைவெளிகள் பெரிதாக இல்லாத சமச்சீரான அல்லது ஓரளவிற்குச் சமச்சீரற்ற பரவல்களில் இந்த ஊகத்தினுல் நிகழும் பிழை குறைந்ததாகவோ அல்லது ஒரு குறிப்பிட்ட குறியை (sign) உடையதாகவோ இருக்கும். மிகவும் சீரற்ற மற்றும் 'J' வடிவப் பரவல்களில் இவ் வூகத்தினுல் தவறு நேர்வதோடு, நிகழ்வெண்கள் யாவும் இடைவெளிப் பிரிவின் ஓர் எல்லையில் குவிவதால், தவறுகள் ஒரு குறிப்பிட்ட குறியை உடையனவாயும் இருக்கும். இம் மாதிரி இடங்களில் பிரிவு இடைவெளிகளை மிகவும் சிறியதாக்கியபின் கணக்கிடப்படும் கூட்டுச் சராசரியின் மதிப்புத் தவறின்றி அமையும். இவ் வூக அமைப்பிற்குப்பின் தொடர்மாறிகளுக்கும் (continuous variable), தனித்த மாறிகளுக்கும் (discrete variable) கணக்கிடும் முறையில் மாற்றமேதுமில்லை.

இடைவெளிப் பிரிவெண்களை (mid-x) $x_1, x_2, \ldots x_n$ எனக் கொண்டு, அவற்றுக்குச் சரியான நிகழ்வெண்களை $f_1, f_2 \ldots f_n$ எனக் கொள்ளும்போது, ஒரு நிகழ்வெண் பரவலுக்கான கூட்டுச் சராசரி

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots + xn\,fn}{f_1 + f_2 + \dots + fn} = \frac{\Sigma\, xi\, fi}{\Sigma\, fi} = \frac{\Sigma\, xi\, fi}{n}$$

என வரையறுக்கப்படும்.

காட்டு :

ஈராயிரத்து முன்னூற்றி முப்பத்துமூன்று நபர்களின் வயதிற்கான கீழ்க்கண்ட நிகழ்வெண் பரவலின் கூட்டுச் சராசரியைக் கணக்கிடுவோம்.

பிரிவு இடை வெளி-வயது	நபர்கள் (நிகழ்வெண்) fi	பிரிவெண் (mid-x) xi	$xi \times fi$
25—26	1	25·5	25·5
27—28	4	27·5	110·0
29—30	165	29·5	4867·5
31—32	357	31·5	11245·5
33—34	169	33·5	5661·5
35—36	682	35·5	24211·0
37—38	196	37·5	7350·0
39—40	381	39·5	15049·5
41—42	80	41·5	3320·0
43—44	48	43·5	2088·0
45—46	162	45·5	7371·0
47—48	57	47·5	3562·5
49—50	13	49·5	643·5
	2333		85505·5

கூட்டுச் சராசரி வயது : $\bar{x} = \frac{\Sigma xi\, fi}{\Sigma fi}$

$$= \frac{85505{\cdot}5}{2333} : 36{\cdot}65$$

சம அளவு இடைவெளிப் பிரிவுகளைக் கொண்ட நிகழ்வெண் பரவல்களில், கணக்கியல் முறையில் மிக்க சுருக்கத்தைக் கையாள முடியும். $x_1\ x_2 \ldots xn$ எனும் இடைவெளிப் பிரிவெண்களிலிருந்து தோராய மதிப்பைக் கழித்து வரும் விலக்கங்களை $(d_1\ d_2 \cdots d_n)$ இடைவெளிப் பிரிவு அளவால் (size of the class interval: c)

பிரிவெண் $(mid-x)$ xi	நிகழ் வெண் fi	$\frac{xi - A}{c}$ ui	$ui\ fi$ +	$ui\ fi$ −
x_1	f_1	u_1	$u_1 f_1$	
x_2	f_2	u_2	$u_2 f_2$	
.	.	.	.	
.	.	.	.	
.	.	.	.	
.	.	.	.	
.	.	.	.	
xn	fn	un	$un\ fn$	
	$\Sigma fi = n$			
			$\Sigma\, ui\, fi$	

வகுத்து வரும் மதிப்புகளை $u_1\ u_2 \ldots u_n$ எனக் குறிப்பிடலாம். $u_1\ u_2 \ldots\ldots u_n$ ஆகியவற்றை அவற்றிற்கொத்த நிகழ்வெண் களால் பெருக்கி வரும் மதிப்புகளின் கூட்டுத்தொகையை, நிகழ் வெண்களின் கூட்டுத் தொகையால் வகுத்து u மாறிக்கான கூட்டுச் சராசரியைக் கணக்கிடுகிறோம்.

$$\bar{u} = \frac{u_1 f_1 + u_2 f_2 + \dots\dots + u_n f_n}{f_1 + f_2 + \dots\dots + f_n} = \frac{\Sigma\, ui\, fi}{\Sigma\, fi} = \frac{\Sigma\, ui\, fi}{n}$$

பிரிவெண் xi	நிகழ்வெண் fi	ui $\frac{xi - 37.5}{2}$	$ui\ fi$	
			+	−
25·5	1	− 6		6
27·5	4	− 5		20
29·5	165	− 4		660
31·5	357	− 3		1071
33·5	169	− 2		338
35·5	682	− 1		682
37·5	196	0	—	
39·5	381	1	381	
41·5	80	2	160	
43·5	48	3	144	
45·5	162	4	648	
47·5	75	5	375	
49·5	13	6	78	
	2333		1786	2777
			− 991	

$$ui = \frac{xi - A}{c}$$

$$xi = A + c\, ui$$

$$\Sigma\, xi = \Sigma\, A + c\, \Sigma\, ui$$

$$\bar{x} = A + c\, \bar{u}$$

$$\bar{u} = \frac{\Sigma\, ui\, fi}{\Sigma fi} = \frac{-991}{\Sigma 333}$$

$$= -0{\cdot}42477.$$

$$\bar{x} = A + c\bar{u} \cdot (A = 37{\cdot}5,\ c : 2).$$

$$= 37{\cdot}5 + 2 \times (-0{\cdot}42477)$$

$$= 37{\cdot}5 - 0{\cdot}8495$$

$$= 36{\cdot}65$$

கூட்டுச் சராசரியின் அடியிற்காணும் கணக்கியல் பண்புகள் புள்ளியியலில் மிகவும் உபயோகமாக அமைந்துள்ளன.

(a) தனிப்பட்ட அளவுகளிலிருந்து கூட்டுச் சராசரியைக் கழித்து அமைக்கும் விலக்கங்களின் கூட்டுத்தொகை பூச்சியத்திற்குச் சமமாகும்.

$$x : \frac{x_1 + x_2 + \ldots\ldots + x_n}{n}$$

இருபக்கமும் கூட்டுச் சராசரியைக் $(\bar{x})$ கழித்தால்

$$\bar{x} - \bar{x} : \frac{x_1 + x_2 + \qquad + x_n}{n} = \bar{x}$$

$$O = \frac{x_1 + x_2 + \ldots\ldots + x_n}{n} - \bar{x}$$

$$O = x_1 + x_2 + \quad\ldots + x_n - n\bar{x}$$

$n\bar{x}$—ஐ எல்லா உறுப்புகளுக்கும் பரப்பும்போது

$$(x_1 - x) + (x_2 - x) + \ldots\ldots (x_n - x) = O$$

தனிப்பட்ட அளவுகளிலிருந்து கூட்டுச் சராசரியைக் கழித்து வரும் விலக்கங்களின் கூட்டுத் தொகை பூச்சியத்திற்குச் சமமாகிறது.

காட்டு :

உயரம் xi	$xi - \bar{x}$ $xi - 66$
59	−7
61	−5
63	−3
65	−1
67	+1
69	+3
71	+5
73	+7
528	0

(b) $\overline{x_1}$ கூட்டுச் சராசரி உள்ள n_1 உறுப்புகளும், $\overline{x_2}$ கூட்டுச் சராசரி உள்ள n_2 உறுப்புகளும் உள்ள இரு மாதிரிகளைச் சேர்த்தால் கிடைக்கும் n $(n_1 + n_2)$ உறுப்புகளின் கூட்டுச் சராசரி

$$\bar{x} = \frac{n_1 \overline{x_1} + n_2 \overline{+} x_2}{n_1 + n_2}$$

என்ற வாய்பாட்டின் மூலம் கிடைக்கும்.

$$\overline{x}_1 : \frac{n_1 \text{ அளவுகளின் கூட்டுத்தொகை}}{n_1}$$

$\therefore$ n_1 அளவுகளின் கூட்டுத்தொகை : $n_1 \overline{x}_1$.

இதேபோல் n_2 அளவுகளின் கூட்டுத்தொகை : $n_2 \overline{x_2}$ என அமையும்.

$(n_1 + n_2)$ அளவுகளின் கூட்டுத்தொகை : $n_1 \overline{x}_1 + n_2 \overline{x}_2$
எனவே n_1, n_2 அளவுகளைச் சேர்ப்பதால் கிடைக்கும் அளவுகளுக்கான கூட்டுச் சராசரி.

$$\overline{x} : \frac{n_1 \overline{x_1} + n_2 \overline{x_2}}{n_1 + n_2}$$

மேற்கண்ட முறைப்படி $n_1, n_2, n_3 \ldots\ldots\ldots nk$ அளவுகளின் கூட்டுச் சராசரிகளாக $\overline{x_1}, \overline{x_2} \ldots\ldots\ldots \overline{x}k$ எனக் கொடுக்கப்பட்ட இடத்து $(n_1 + n_2 + \ldots + nk)$ அளவுகளின் கூட்டுச் சராசரி

$$\bar{x} \cdot = \frac{(n_1 \bar{x}_1 + n_2 \bar{x}_2 + \ldots + nk \, \overline{x}k)}{n_1 + n_2 + \ldots + nk}$$

என அமையும்.

காட்டு :

$2333 \ (n_1)$ நபர்களின் கூட்டுச் சராசரி வயது $86 \cdot 65$ எனவும் $1078 \ (n_2)$ நபர்களின் கூட்டுச் சராசரி வயது $36 \cdot 24$ எனவும் கொடுக்கப்பட்டுள்ளன. இவ் விரு மாதிரிகளையும் சேர்க்கும்போது கிடைக்கும் கூட்டுச் சராசரியின் மதிப்பைக் கணக்கிடுவோம்.

கொடுக்கப்பட்ட அளவுகள் :

$n_1 : 2333.$ $\quad \overline{x}_1 : 36 \cdot 65$

$n_2 : 1078.$ $\quad \overline{x}_2 : 36 \cdot 24$

$$\overline{x} : \frac{n_1 \overline{x}_1 + n_2 \overline{x}_2}{n_1 + n_2}$$

$$: \frac{(36{\cdot}65 \times 2333) + (1078 \times 36{\cdot}24)}{2333 + 1078}$$

$$: \frac{124571{\cdot}17}{3411} = 36{\cdot}52$$

இடைநிலை (median)

தொகுக்கப்பட்ட விவரங்களை அவற்றின் மதிப்பிற்குத் தக்கவாறு ஏறு அல்லது இறங்கு வரிசையில் அமைத்த பின் கிடைக்கும் வரிசையின் (array) மத்தியில் அமைத்து அதற்கு மேலும் கீழும் சமமான உறுப்புகளைக் கொள்ளும் மதிப்பு இடைநிலை (median) எனப்படும். இது தொடர்வரிசையின் மத்தியில் அமைந்து இருபுறங்களிலும் சமஎண்ணிக்கை உறுப்புகளைக் கொண்டதாக இருக்கும்.

இடைநிலை மதிப்பு மையத்தில் அமைவதால், வரிசைப் படுத்தப்பட்ட 'n' உறுப்புகளில் $\left(\frac{n+1}{2}\right)$-வது உறுப்பு இடைநிலை மதிப்பாக அமையும். உறுப்புகளின் மொத்த எண்ணிக்கை ஒற்றைப் படை எண்ணாக இருப்பின், $\left(\frac{n+1}{2}\right)$-வது உறுப்பு வரிசையில் அமைந்திருப்பதை எளிதாக உணரமுடியும். $x_1, x_2, \ldots\ldots\ldots$ x_{2n+1} என்பன அமைக்கப்பட்ட வரிசையில் உள்ள உறுப்புகளாயின் $\frac{(2n+1)+1}{2}$: $\frac{2n+2}{2}$: $(n+1)$ வது உறுப்பு, அதாவது x_{n+1} என்ற உறுப்பு இடைநிலை மதிப்பாக அமையும். $\left(\frac{n+1}{2}\right)$ என்ற அளவு இடைநிலை மதிப்பைக் குறிப்பிடாமல், இடைநிலை மதிப்பிற்கான உறுப்பைக் கொடுக்கும் வாய்பாடாக உள்ளதை மனதில் கொள்ளுதல் நலம். அதாவது $\left(\frac{n+1}{2}\right)$ க்குச் சமமான உறுப்பினை வரிசையிலிருந்து எடுத்து இடைநிலை மதிப்பைக் குறிப்பிட வேண்டும்

காட்டு :

கொடுக்கப்பட்ட அளவுகள் : 5 7 9 12 10 8 7 15 21
(raw data)

வரிசைப் படுத்தப்பட்டவை : 5 7 7 8 (9) 10 12 15 21
(array)

மொத்த உறுப்புகளின் எண்ணிக்கை : 9. (ஒற்றைப்படை எண்).

இடைநிலை வாய்பாடு : $\frac{n+1}{2}$: $\frac{9+1}{2} = 5$ வது உறுப்பு

இடைநிலை : $x_5 = 9$

இடைநிலை மதிப்பான ஒன்பதிற்கு மேலும் கீழும் சமமான உறுப்புகள் அமைந்திருப்பதைக் காணலாம்.

கொடுக்கப்பட்ட அளவுகளின் மொத்த எண்ணிக்கை இரட்டைப்படை எண்ணாக அமையும்போது, இடைநிலை மதிப்பு கொடுக்கப்பட்ட ஓர் உறுப்பாக அமையாமல், இரு உறுப்பு களுக்கு மத்தியில் அமையும் ஒரு மதிப்பீடாகக் கொள்ளப்படும். ஏறு அல்லது இறங்கு வரிசையில் உறுப்புகள் $x_1\ x_2 \ldots\ x_n$ என அமையும்போது, இடைநிலை $\left(-\frac{2n}{2}\right)$ மற்றும் $\left(\frac{2n}{2}+1\right)$-வது $\left[\left(-\frac{n}{2}, \frac{n}{2}+1\right)\right]$ ஆகிய இரு உறுப்புகளின் மத்தியில் அமை கிறது. இவ்விரு உறுப்புகளின் கூட்டுச் சராசரியே இடை நிலை மதிப்பெனக் கொள்கிறோம்.

இடைநிலை :

$$\frac{\left(\frac{2n}{2}\right)\text{-வது உறுப்பின் மதிப்பு} + \left(-\frac{2n}{2}+1\right)\text{-வது உறுப்பின் மதிப்பு}}{2}$$

$$= \frac{x_n + x_{n+1}}{2}$$

காட்டு :

கொடுக்கப்பட்ட அளவுகள் :	45	25	29	28	33	40
(raw data)	30	38	32	35	34	42
வரிசைப் படுத்தப்பட்டவை:	25	28	29	30	32	(33)
(array)	(34)	35	38	40	42	45.

மொத்த உறுப்புகள் : 12. இரட்டைப்படை எண். (2n)

இடைநிலை வாய்ப்பாடு :

$$\frac{\left(\frac{2n}{2}\right)\text{-வது உறுப்பு மதிப்பு} + \left(\frac{2n}{2}+1\right)\text{-வது உறுப்பு மதிப்பு}}{2}$$

$$= \frac{6\text{ வது உறுப்பு} + 7\text{ வது உறுப்பு}}{2}$$

$$இடைநிலை = \frac{33 + 34}{2} = 33 \cdot 5$$

இடைநிலைக்கு இருபுறமும் சம எண்ணிக்கை உறுப்புகள் அமைந்திருப்பதைக் காணலாம்.

இடைநிலை : தனித்த மாறிக்கானது (Median : Descrete variable)

தனித்த மாறிகளுக்கு இடைநிலை மதிப்பைக் கணக்கிடுவது அவற்றின் அமைப்பு முறையில் மிக எளிதாக உள்ளது.

குடும்பத்திலுள்ள குழந்தைகளின் எண்ணிக்கை	குடும்பங்களின் எண்ணிக்கை	குடும்பத்திலுள்ள குழந்தைகளின் எண்ணிக்கை	குடும்பங்களின் எண்ணிக்கை
0	96	7	20
1	108	8	11
2	154	9	6
	126	10	5
4	95	11	5
5	62	12	1
6	45	13	1
			735.

குடும்பத்திலுள்ள குழந்தைகளின் எண்ணிக்கைக்குத் தக்கபடி இத் தனித்த மாறியின் அளவுகள் ஏறுவரிசையில் (acending order) அமைந்திருப்பதைக் காண்கிறோம். குழந்தைகளே இல்லாத குடும்பம் 96 எனவும் ஒரு குழந்தை உள்ள குடும்பங்கள் 108 எனவும் ....பதின்மூன்று குழந்தைகள் உள்ள குடும்பம் 1எனவும் குறிக்கப்பட்டுள்ளது. இடைநிலை மதிப்பானது இருபுறமும் சமமான குடும்பங்கள் அமையப்பெற்ற குடும்பத்தில் உள்ள குழந்தைகளின் எண்ணிக்கையைக் குறிக்கும்.

குழந்தைகளின் எண்ணிக்கை x	வளர் நிகழ்வெண் $<x$		குழந்தைகளின் எண்ணிக்கை	வளர் நிகழ்வெண்
0	96	$x_1\ x_2 \ldots\ldots x_{96}$	7	706
1	204	$x_{96} \ldots\ldots x_{204}$	8	717
2	358	$x_{205} \ldots x_{358}$	9	723
3	484	$x_{309} .. x_{368}\quad x_{484}$	10	728
4	579	$x_{485} \ldots\ldots x_{579}$	11	733
5	641	etc	12	734
6	686		13	735

வரிசைப்படி அமைக்கும்போது 368 ஆவது குடும்பத்திலுள்ள $\left(\frac{735+1}{2} : 368\right)$ குழந்தைகளின் எண்ணிக்கையே இடைநிலை மதிப்பாகும். வரிசைப்படுத்தி அமைக்கும் நிலை இங்கு வளர் நிகழ்வெண்பரவலைச் (cumulative frequency distribution) சுட்டும். வளர் நிகழ்வெண் பரவலில் 368 வது குடும்பம் 3 குழந்தைகள் உள்ள குடும்பத்தைச் சுட்டுகிறது.

மொத்தக் குடும்பங்களின் எண்ணிக்கை : 735.

இடைநிலை மதிப்பிற்கான குடும்பம் : $\frac{735+1}{2}$: 368.

: x_{368}

இடைநிலை மதிப்பு : 3.

இடைநிலை மதிப்பு : நிகழ்வெண் பரவலுக்கானது (Median : Frequency distribution).

தொடர் மாறியில், உறுப்புகளின் மதிப்புகள் தெரியாத நிலையாய் உள்ளதால், இடைநிலை மதிப்புச் சார்ந்துள்ள இடைவெளியைக் கணக்கிட்டு, அடியிற்காணும் வாய்பாட்டின் உதவியால் இடைநிலை மதிப்பைக் கணக்கிடுகிறோம்.

இடைநிலை : $m : Q_2 : \; l + \left(\frac{n/2 - m}{f}\right) C.$

l : இடைநிலை இடைவெளிப்பிரிவின் கீழ் எல்லை மதிப்பு

m : இடைநிலை இடைவெளிப்பிரிவின் கீழ் எல்லை வரை அமையும் கீழின வளர் நிகழ்வெண்களின் மதிப்பு.

f : இடைநிலை இடைவெளிப் பிரிவின் நிகழ்வெண்.

c : இடைநிலை இடைவெளிப் பிரிவின் அளவு.

கொடுக்கப்பட்ட மதிப்புகள் x_1 x_2 . ... x_n எனக் கொள்ளும் போது $x_n/2$ உறுப்பினைக் கொண்டுள்ள இடைவெளிப்பிரிவு இடைநிலை இடைவெளிப் பிரிவாகும் (Medianal class). தனித்த மாறிகளுக்காக அமைக்கப்பட்ட உள்ளடக்கிய இடைவெளிப் பிரிவுகளின் உண்மை இடைவெளிப் பிரிவுகளை அமைத்து இடைநிலை மதிப்பைக் கணக்கிடவேண்டும்.

காட்டு :

இடைவெளிப் பிரிவு	நிகழ்வெண்	கீழின வளர் நிகழ்வெண்	உண்மை இடைவெளி
20 – 21	1	1	19·5 – 21·5
22 – 23	13	14	21·5 – 23·5
24 – 25	433	447	23·5 – 25·5
26 – 27	496	943	25·5 – 27·5
28 – 29	398	1341	27·5 – 29·5
30 – 31	665	2006	29·5 – 31·5
32 – 33	162	2168	31·5 – 33·5
34 – 35	103	2271	33·5 – 35·5
36 – 37	45	2316	35·5 – 37·5
38 – 39	11	2327	37·5 – 39.5
40 – 41	6	2333	39·5 – 41·5
	2333		

$\frac{n}{2}$ $\frac{2333}{2}$: 1166 · 5, x_{1167}-வது உறுப்பு அமையும் இடைவெளி, இடைநிலை இடைவெளி எனப்படும். கீழின வளர் நிகழ் வெண்ணை நோக்குங்கால் x_{1167}-வது உறுப்பு அமையும் இடைவெளி (28 – 29) இடைநிலை இடைவெளியாகும்.

l : 27·5 (உண்மை இடைவெளிப் பிரிவின் கீழ் எல்லை)

m : 943

f : 398

c : 2

$\frac{N}{2}$: 1166·5

இடைநிலை : m : $Q_2 \; l + \left(\frac{N/2 - m}{f}\right) c$

$$: 27·5 + \left(\frac{1166·5 - 943}{398}\right) ·2$$

$$: 27·5 + 1·123$$

$$: 28·623$$

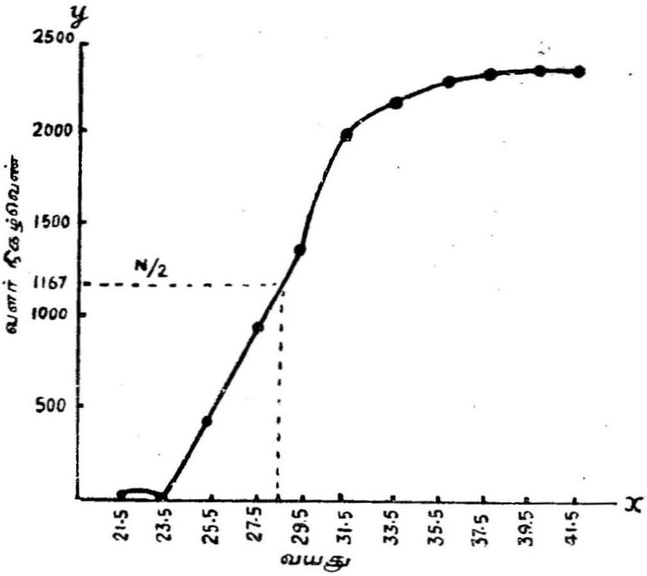
கணக்கியல் குறியீடாக இடைநிலை மதிப்பு 'M' அல்லது Q_2 எனக் குறிக்கப்படும்.

வரைபட மூலம் இடைநிலை மதிப்பைக் கணக்கிடுதல் :
(location of median by Grahpical method)

கீழினம் அல்லது மேலின வளர் நிகழ்வெண்களின் வளை கோட்டினை உபயோகப்படுத்தி, இடைநிலையைக் கணக்கிடலாம். y-அச்சில் $\frac{N}{2}$-க்குச் சமமாக, x அச்சிற்கு இணைகோடொன்று வரைந்து, இவ் விணைகோடு வளைகோட்டைத் தொடும் இடத்திலிருந்து, x அச்சுக்கு வரையப்படும் செங்குத்துக்கோடு x-அச்சைத் தொடும் இடத்திற்கான x-மாறியின் மதிப்பே இடைநிலையின் மதிப்பாகும். கீழினம் மற்றும் மேலின வளர் நிகழ்வெண் பரவல்களுக்கான வளைகோடுகள் வரையும்போது இவ்விரு வளைகோடுகளும் ஒன்றையொன்று வெட்டுமிடத்திலிருந்து x-அச்சுக்கு வரையப்படும் செங்குத்துக்கோடு x–அச்சைத்தொடும் இடத்திற்கான x-மாறியின் மதிப்பும் இடைநிலை மதிப்பிற்குச் சமமாக அமைவதைக் காணலாம்.

காட்டு :

மேலே குறிப்பிட்ட வயதிற்கான நிகழ்வெண் பரவலின் இடைநிலை மதிப்பை வரைபட மூலம் மதிப்பிடுவோம்.



படம் 3·1

இடைநிலை : M : 28·6

முகடு (Mode) :

கொடுக்கப்பட்ட நிகழ்வெண் பரவல் வளைகோட்டின் உச்ச மதிப்பே முகடாகும். அடிக்கடி நிகழும் அல்லது பெரும்பாலாக அமையும் அளவினையே இது பிரதிபலிக்கவேண்டும். அளவுகளின் உயர் மதிப்பை ஏற்பதோடு, இவ் வுயர்மதிப்பைச் சுற்றி நிகழ வெண்களின் அடர்த்தியும் அமையப்பெற்றதாய் இருக்க வேண்டும். முகட்டின் முன்னும் பின்னும் அமைந்துள்ள நிகழ் வெண்களின் குவிவு வேறெந்த மதிப்பிற்குமில்லாதபடி அதிக அளவில் குவிந்திருக்க வேண்டும். நிகழ்வெண்களின் குவிவு (concentration) இல்லாத உச்ச மதிப்பைக் காட்டிலும், சற்றுக் குறைந்த மதிப்புடைய ஆனால் நிகழ்வெண்களின் குவிவு உள்ள மதிப்பே முகடாகக் கொள்ளப்படும்.

நிகழ்வெண் பரவலில் அதிக நிகழ்வெண் உள்ள இடைவெளிப் பிரிவை முகட்டிடை வெளிப்பிரிவாகக் (model class) கொண்டு, அடியிற்காணும் வாய்பாட்டின் மூலம் முகட்டின் மதிப்புக் கணக் கிடப்படுகிறது.

$$\text{முகடு}: \quad Z : l + \left(\frac{f - f_1}{2f - f_1 - f_2}\right) c.$$

l : முகட்டிடை வெளிப் பிரிவின் கீழ் எல்லை மதிப்பு

f : முகட்டிடை வெளிப் பிரிவு நிகழ்வெண்

f_1 : முகட்டிடை வெளிக்கு முந்திய இடைவெளிப் பிரிவு நிகழ்வெண்

f_2 : முகட்டிடை வெளிக்குப் பிந்திய இடைவெளிப் பிரிவு நிகழ்வெண்

c : பிரிவு இடைவெளி அளவு

அதிக நிகழ்வெண் உள்ள இடைவெளியே பொதுவாக முகட்டிடை வெளியாகக் கொள்ளப்படும். முகட்டின் இரண்டாவது பண்பாகிய இருபுறமும் அதிக அடர்த்தி இருக்க வேண்டுமென்ற பண்பின் அடிப்படையில், முகட்டிடை வெளியைக் கணக்கிடச் சிறு மாற்றுமுறைகளைக் கையாளுவதும் உண்டு. நமது தேவைக்கு அதிக நிகழ்வெண் உள்ள இடைவெளிப்பிரிவே முகட்டிடை வெளியாகக் கொள்ளப்படுவது போதுமானதாகும்.

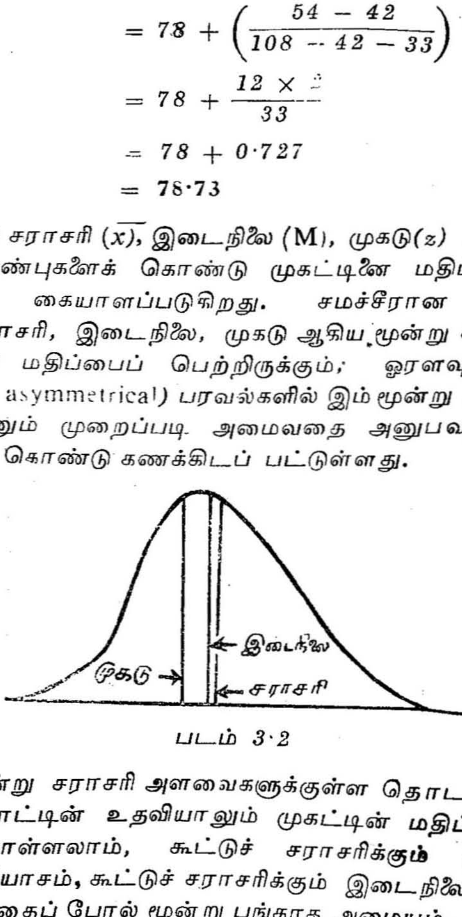
காட்டு :

இடைவெளிப் பிரிவு	நிகழ்வெண்	
72 – 74	7	
74 – 76	31	
76 – 78	42	
78 – 80	54	.....முகட்டிடைவெளி (Model class)
80 – 82	33	
82 – 84	24	
84 – 86	22	
86 – 88	8	
88 – 90	4	
	225	

முகட்டிடை வெளி : 78—80.

$$l : 78$$
$$f : 54., f_1 : 42, f_2 : 33.$$
$$C : 2.$$
$$Z : l + \left(\frac{f - f_1}{2f - f_1 - f_2}\right) C$$
$$= 78 + \left(\frac{54 - 42}{108 - 42 - 33}\right) \times 2$$
$$= 78 + \frac{12 \times 2}{33}$$
$$= 78 + 0{\cdot}727$$
$$= 78{\cdot}73$$

கூட்டுச் சராசரி ($\bar{x}$), இடைநிலை (M), முகடு(z) ஆகியவற்றின் பொதுப் பண்புகளைக் கொண்டு முகட்டினை மதிப்பிடுவது சில இடங்களில் கையாளப்படுகிறது. சமச்சீரான பரவல்களில் கூட்டுச் சராசரி, இடைநிலை, முகடு ஆகிய மூன்று சராசரி அளவுகளும் ஒரே மதிப்பைப் பெற்றிருக்கும்; ஓரளவு சமச்சீரற்ற (moderately asymmetrical) பரவல்களில் இம் மூன்று அளவைகளும் அடியிற்காணும் முறைப்படி அமைவதை அனுபவத்தை அடிப்படையாகக் கொண்டு கணக்கிடப் பட்டுள்ளது.



படம் 3·2

இம் மூன்று சராசரி அளவைகளுக்குள்ள தொடர்பை இணைக்கும் வாய்பாட்டின் உதவியாலும் முகட்டின் மதிப்பைக் கணக்கிட்டுக் கொள்ளலாம், கூட்டுச் சராசரிக்கும் முகட்டிற்கும் உள்ள வித்தியாசம், கூட்டுச் சராசரிக்கும் இடைநிலைக்கும் உள்ள வித்தியாசத்தைப் போல் மூன்று பங்காக அமையும். நிகழ்வெண் பரவல் சம அளவு இல்லாத இடைவெளிகளைப் (unequal class intervals) பெற்றும், பரவலின் தன்மை ஓரளவிற்கு மட்டும் சமச்சீரற்ற நிலையில் உள்ள போதும், முகட்டின் மதிப்பை இம் முறைப்படி கணக்கிடுவது நடைமுறையில் உள்ளது.

(கூட்டுச் சராசரி — முகடு) = 3(கூட்டுச் சராசரி — இடைநிலை)

$$(\overline{x} - z) = 3\,(\overline{x} - m)$$

காட்டு :

அடியிற்காணும் நிகழ்வெண் பரவலுக்கான முகட்டினை (mode) வரையறுக்கப்பட்ட வாய்பாட்டிலிருந்தும், மேலே கொடுக்கப்பட்ட வாய்பாட்டிலிருந்தும் (கூட்டுச் சராசரி, இடைநிலை மதிப்புகளிலிருந்து கணக்கிடுவது) கணக்கிடுவோம்.

இடை வெளிப் பிரிவு	70—80	80—90	90—100	100—110	110—120	120—130	130—140	140—150
நிகழ் வெண்	12	18	35	42	50	45	20	8

இடைவெளிப் பிரிவு	நிகழ் வெண்	பிரி வெண்	u_i	$u_i f_i$	கீழின வளர் நிகழ்வெண்
70—80	12	75	−4	−48	12
80—90	18	85	−3	−54	30
90—100	35	95	−2	−70	65
100—110	42	105	−1	−42	107
				−214	
110—120	50	115	0		157
120—130	45	125	1	45	202
130—140	20	135	2	40	222
140—150	8	145	3	24	230
				+109	
	230			−105	

$$\overline{x} = 115 + 10\left(-\frac{105}{230}\right)$$
$$= 115 - 4565$$
$$= 110{\cdot}435$$

$$M = l + \left(\frac{N/2 - m}{f}\right)C.$$
$$= 110 + \left(\frac{115 - 117}{50}\right) 10$$
$$= 110 + \frac{8}{5} = 111{\cdot}6$$

$$Z = l + \left(\frac{t - t_1}{2t - t_1 - t_2}\right) C$$
$$= 110 + \left(\frac{50 - 42}{100 - 42 - 45}\right) 10$$
$$= 110 \div 6{\cdot}156 = 116{\cdot}156$$

முகடு = (கூட்டுச் சராசரி — முகடு) = 3 (கூட்டுச் சராசரி — இடைநிலை)

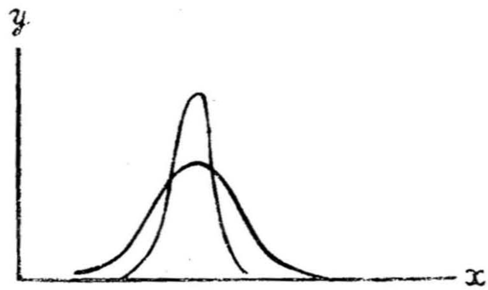
முகடு $= 110{\cdot}435 - 3\ (110{\cdot}435 - 111{\cdot}6)$
$= 110{\cdot}435 + 3{\cdot}495$
$= 113{\cdot}930$

கூட்டுச் சராசரி எளிதாகக் கணக்கிடப்படுவதாலும், எல்லாச் சமயங்களிலும் ஒரே மாதிரியான அளவாய் அமைவதாலும், எல்லா உறுப்புகளையும் கொண்டு மதிப்பிடப்படுவதாலும் மாதிரிகளின் ஏற்றத் தாழ்வுகளினால் (Sampling fluctuations) அதிகம் பாதிக்கப்படாததாலும், கணக்கியல் முறைக்கு இலக்காக இருப்பதாலும், புள்ளியியலில் மற்றச் சராசரி அளவைகளைக் காட்டிலும் இதன் உபயோகம் அதிகமாக அமைந்துள்ளது. மாதிரியின் ஏற்றத் தாழ்வுகளினால் பாதிக்கப்படுவதாய், எல்லா உறுப்புகளையும் உட்படுத்தாமல் கணக்கிடப்படுவதாய், சில சமயங்களில் சரியாகக் கூற முடியாத தன்மைத்தான இடைநிலை (Median), கூட்டுச் சராசரியை மிகவும் பாதிக்கும் எல்லை உறுப்புகள் இருக்குமிடத்து ஒரு சிறந்த சராசரி அளவையாக அமைகிறது. சில மாறிகளின் தன்மையை விவரிக்கும்போது கூட்டுச் சராசரி அளவையில் சொல்லாமல் இடைநிலை அளவையில் குறிப்பிடுவது மரபாக உள்ளது. சராசரி அறிவுள்ள மாணவன் அல்லது சராசரி அழகுள்ள பெண் எனக் குறிப்பிடுவதைக் காட்டிலும் இடைநிலை அறிவு, இடைநிலை அழகு எனக் குறிப்பிடுவது பொருளுடையதாய் அமைகிறது. முகட்டின் (mode) நிலையிலாத் தன்மை அதனை அதிக அளவில் உபயோகப்படுத்துவதைத் தடுத்தாலும், மாறி

களின் தன்மைக்கேற்பச் சில இடங்களில் சராசரிஅளவு முகட்டால் குறிக்கப்படும். குடும்பங்களின் குழந்தைகளுக்கான கூட்டுச் சராசரி அளவு இவ்வளவு என்பதைக் காட்டிலும், பெரும்பாலான குடும்பங்களின் (அடிக்கடி நிகழும் தன்மைத்தான) குழந்தைகளின் எண்ணிக்கையைக் (முகடு) குறிப்பிடுவது சிறந்ததாகும். கொடுக் கப்பட்ட மாறியின் தன்மைக்கேற்ற சராசரி அளவைகளைத் தேர்ந்தெடுப்பது அனுபவத்தால் உணரக் கூடிய ஒன்றாகும்.

சிதறல் அளவைகள் (Measures of dispersion)

சராசரிகளின் உதவியால் இரு பரவல்களை ஒப்பிடும்போது தவறான முடிவுகளுக்கு வருவதைத் தவிர்க்க முடியாதவர்களாக இருக்கிறோம். மூல அளவுகள் குறிப்பிடப்படாமல் சராசரி அளவைகள் மட்டும் குறிப்பிடும் இடங்களில் இதன் விளைவு மிகவும் அதிகமாக்கப்படும். ஒரு மாறியின் அளவு 40, 50, 60, 70, 80 எனவும் 58, 55, 65, 60, 62 எனவும் உள்ள இரு இடங் களிலும் 60 என்ற கூட்டுச் சராசரியையே கொடுக்கின்றன. சராசரிகளின் மதிப்புகள் ஒன்றாக இருப்பினும் மாறிகளின் தனிப் பட்ட அளவுகள் இச் சராசரியைச் சுற்றி அமைந்துள்ள முறையில் மாற்றம் இருப்பதை உணர்கிறோம்.



படம் 3·3

ஒரு மாறியின் அளவுகளைப் படிக்கும்போது, அவற்றின் பிரதிநிதித்துவ மதிப்பாகச் சராசரிகளைப் படிப்பதோடு, அச் சராசரியைச் சுற்றிலும் மாறியின் அளவுகள் கூட்டமாக அமைந் துள்ளனவா (clustering around) அல்லது சராசரியினின்றும் சிதறி உள்ளனவா (scattered away) என்பதைக் காணுவதும் இன்றியமை யாததாகும். இத் தன்மையைக் குறிக்கும் அளவைகளே சிதறல் அளவைகள் (measures of dispersion) எனப்படும். இவ் வளவை களே சிதறல், விலகல், வேறுபாடு, மாறுபாட்டளவைகள் எனப் பல பெயரால் கூறப்படுகின்றன.

திட்ட விலக்கம் (Standard deviation)

மாறியின் அளவுகளிலிருந்து கூட்டுச் சராசரியினைக் கழித்து வரும் விலக்க வர்க்கங்களின் கூட்டுத் தொகையை, மாறியின் மொத்த எண்ணிக்கையால் வகுத்து வரும் மதிப்பின் வர்க்க மூலமே (Square root) திட்ட விலக்கம் எனப்படும். மாதிரியின் (sample) திட்ட விலக்கத்தை 's' எனவும், முழுமைத் தொகுதியின் (population) திட்ட விலக்கத்தை 'σ' எனவும் குறிப்பிடுவது பழக்கத்தில் உள்ளது. முழுமைத் தொகுதியின் மற்றும் மாதிரியின் மாறுபட்டவை (Variance) கணக்கிடப்பட்டு, அவற்றின் வர்க்க மூலமாகத் திட்ட விலக்க மதிப்புகள் குறிக்கப்படும்.

$$s : \sqrt{\frac{1}{n} \Sigma fi\,(xi - \bar{x})^2}$$

விலக்கங்களின் வர்க்கத்தைக் காணுவது செயற்கையாகப் பட்டாலும், மாறியின் அளவுகளிலிருந்து கூட்டுச்சராசரியைக் கழித்து அமைக்கப்படும் விலக்கங்களின் கூட்டுத்தொகை பூச்சியமாக முடிவதால், இது ஒரு சிறந்த முறையாகவே அமைந்துள்ளது. கூட்டுச்சராசரியினின்றும் எடுக்கப்படும் விலக்கங்களில் சில கூட்டெண்ணுகவும் (positive number) சில குறையெண்ணுகவும் (negative number) இருப்பதால் இவற்றின் கூட்டுத் தொகை பூச்சியமாக அமைகிறது. இவற்றை ஒரே குறி (Sign) உடையதாக்குவதற்கு, இவற்றின் வர்க்கங்களைக் காணுவது சிறப்பாக அமைந்துள்ளது.

கொடுக்கப்பட்ட அளவுகளை x_1 x_2 ... x_n எனவும் கூட்டுச் சராசரியை $\bar{x}$ எனவும் குறிப்பிட்டு $(x_1 - x)$, $(x_2 - \bar{x})$......$(x_n - \bar{x})$ ஆகிய விலக்கங்களை அமைத்துத் திட்டவிலக்கம் கணக்கிடப்படுகிறது.

xi	$(xi - \bar{x})$	$(xi - \bar{x})^2$
x_1	$x_1 - \bar{x}$	$(x_1 - x)^2$
x_2	$x_2 - \bar{x}$	$(x_2 - \bar{x})^2$
.	.	.
.	.	.
.	.	.
x_n	$(x_n - \bar{x})$	$(x_n - x)^2$
Σx_i		$\Sigma(xi - \bar{x})^2$

$$\bar{x} : \frac{\Sigma xi}{n}.$$

$$s : \sqrt{\frac{\Sigma (xi - \bar{x})^2}{n}}$$

காட்டு : மாறியின் அளவுகள் 145, 148, 150, 154, 155, 155. 156, 159, 162 என அமைந்த இடத்துக் கிடைக்கும் திட்ட விலக்கத்தைக் (s : standard deviation S.D.) கணக்கிடுவோம்.

x_i	$(x_i - \bar{x})$	$(x_i - x)^2$
145	−9	81
148	−6	36
150	−4	16
154	0	0
155	1	1
155	1	1
156	2	4
156	2	4
159	5	25
162	8	64
1540		232

$$\bar{x} : \frac{1540}{10} = 154.$$

$$S.D : s : \sqrt{\frac{\Sigma (x_i - x)^2}{n}}$$

$$: \sqrt{\frac{232}{10}} = \sqrt{23 \cdot 2}$$

$$\mathbf{S \cdot D} = 4 \cdot 817$$

விலக்கங்களை அமைக்காமலேயே திட்ட விலக்கத்தைக் காண்பதற்கு ஓர் எளிய முறை உள்ளது. பின்னமாக உள்ள விலக்கங்களை வர்க்கமாக்கும்போது ஏற்படும் சிரமம், காலதாமதம் ஆகியவற்றைத் தவிர்க்க இது உபயோகப்படுகிறது.

$$s^2 = \frac{1}{n} \Sigma (x_i - \bar{x})^2$$

$$= \frac{1}{n} \Sigma (x_i^2 - 2 x_i\bar{x} + \bar{x}^2)$$

$$= \frac{1}{n} \Sigma x_i^2 - 2\bar{x} \frac{1}{n} \Sigma x_i + \frac{1}{n} \Sigma \bar{x}^2$$

$$= \frac{1}{n} \Sigma x_i^2 - 2\bar{x}^2 + \bar{x}^2$$

$$= \frac{1}{n} \Sigma x_i^2 - \bar{x}^2$$

மேற்கண்ட கணக்கிற்கான திட்ட விலக்கம் இம் முறைப்படி கணக்கிடப்படும்போது

$$s^2 = \frac{1}{n} (x_1^2 + x_2^2 + \dots + x_n^2) - \bar{x}^2$$

$$= \frac{1}{10} [145^2 + 148^2 + \dots + 162^2] - (154)^2$$

$$= 23 \cdot 2$$

$$s = \sqrt{23 \cdot 2}$$

$$= 4 \cdot 817$$

திட்டவிலக்கம் : தனித்த மாறிக்கானது
(Standard Deviation : Discrete variable)

மாறியின் அளவுகள் ஒரு முறைக்குமேல் நிகழும் இடங்களில் நிகழ்வெண் பரவலை அமைத்து, x_1 என்ற அளவு f_1 முறை நிகழ்வதாகவும், x_2 என்ற அளவு f_2 முறை நிகழ்வதாகவும் .... x_n என்ற அளவு f_n முறை நிகழ்வதாகவும் கொண்டு திட்ட விலக்கம் கணக்கிடப்படுகிறது.

x_i	f_i	$x_i f_i$	$x_i^2 f_i$
x_1	f_1	$x_1 f_1$	$x_1^2 f_1$
x_2	f_2	$x_2 f_2$	$x_2^2 f_2$
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
x_n	f_n	$x_n f_n$	$x_n^2 f_n$
	Σf_i	$\Sigma x_i f_i$	$\Sigma x^2 f_i$

$$s^2 = \frac{1}{n} \Sigma x_i^2 f_i - \bar{x}^2$$

$$= \frac{1}{n} \Sigma x_i^2 f_i - \left(\frac{\Sigma x_i f_i}{n}\right)^2$$

$$s : \sqrt{\frac{1}{n} \Sigma x_i^2 f_i - \left(\frac{\Sigma x_i f_i}{n}\right)^2}$$

காட்டு :

உயரம் அங். x	நபர்கள் f_i	$x_i f_i$	$x_i^2 f_i$
63	2	126	7938
64	1	64	4096
65	1	65	4225
66	2	132	8712
68	3	204	13872
69	2	138	9522
	11	729	48365

$$s^2 : \frac{1}{n} \Sigma x_i^2 f_i - \left(\frac{\Sigma x_i f_i}{n}\right)^2$$
$$= \frac{48365}{11} - \left(\frac{729}{11}\right)^2$$
$$= 4{\cdot}7473$$
$$s = \sqrt{4{\cdot}7473} = 2{\cdot}179$$

திட்டவிலக்கம் : நிகழ்வெண் பரவலுக்கானது
(Standard Deviation : Frequency distribution)

நிகழ்வெண் பரவலுக்கான திட்டவிலக்கம் காணும்போது பிரிவெண்களை (mid-x) $x_1\, x_2 \ldots\ldots x_n$ எனக் கொண்டு மேற் சொன்ன முறைப்படியோ அல்லது, $x_1\, x_2 \ldots\ldots x_n$ எனும் மாறியை u மாறியாக மாற்றிச் சுலபமான முறையிலோ கணக்கிடலாம்.

காட்டு : நேர்முறை.

இடைவெளிப் பிரிவு	நிகழ்வெண் f_i	பிரிவெண் x_i	$x_i f_i$	$x_i^2 f_n$
25—28	8	26·5	212·0	5618·00
29—32	259	30·5	7899·5	240934·75
32—36	367	34·5	12661·5	436821·75
37—40	258	38·5	9933·0	382420·50
41—44	72	42·5	3060·0	130050·00
45—48	114	46·5	5301·0	246496.50
	1078	—	39067·0	1442341·50

$$\bar{x} : \frac{39067}{1078} = 36{\cdot}24$$

$$s^2 = \frac{1}{n} \Sigma x^2 f_i - \bar{x}^2$$
$$= \frac{1442341{\cdot}50}{1078} - (36{\cdot}24)^2$$
$$= 1337{\cdot}9791 - 1313{\cdot}3376$$

$$= 24{\cdot}64$$

$$s_x = \sqrt{24{\cdot}64} = 4{\cdot}964$$

திட்டவிலக்கம் : சுலப முறை

$$u_i = \frac{x_i - A}{c}$$

$$x_i = A + cu_i.$$

$$\bar{x} = A + c\bar{u}.$$

$$S_X^2 = \frac{1}{n} \Sigma f_i (x_i - \bar{x})^2$$

$$\frac{1}{n} \Sigma f_i (A + cu_i - A - c\bar{u})^2$$

$$= \frac{1}{n} \Sigma f_i (c u_i - c\bar{u})$$

$$= \frac{1}{n} \Sigma f_i \cdot c^2 (u_i - \bar{u})^2$$

$$= \frac{1}{n} \Sigma f_i(u_i - \bar{u})^2 \cdot c^2$$

$$= c^2 \cdot S^2_u$$

$$S_X^2 = c^2 \cdot S^2_u.$$

$$S_X = c \cdot S_u$$

இடை வெளிப் பிரிவு	நிகழ் வெண்	பிரி வெண்	u_i $\frac{x_i - 34.5}{4}$	$u_i f_i$	$u_i^2 f_i$
25 – 28	8	26·5	−2	−16	32
29 – 32	259	30·5	−1	−259	259
32 – 36	367	34·5	0	+ 0	·
37 – 40	258	38·5	1	+258	258
41 – 44	72	42·5	2	+144	288
45 – 48	114	44·5	3	+342	1026
	1078			+469	1863

$$\bar{u} : \frac{469}{1078} = 0{\cdot}435.$$

$$s_u^2 : \frac{1863}{1078} - (0{\cdot}435)^2$$

$$: (1{\cdot}7282 - 0{\cdot}189225)$$

$$: 1{\cdot}539.$$

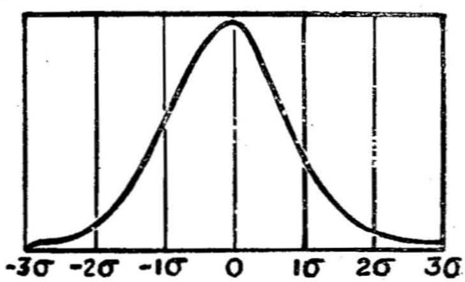
$$s_u : \sqrt{1{\cdot}539} : 1{\cdot}241.$$

$$s_x : c \cdot s_u = 4 \times 1{\cdot}241$$

$$: 4{\cdot}964.$$

சிதறலை அளப்பதற்குள்ள அளவைகளில் திட்டவிலக்கமே மிகவும் சிறந்ததாக அமைந்துள்ளது. சராசரிகளில் கூட்டுச் சராசரியின் உபயோகத்திற்கொப்பச் சிதறல் அளவைகளில் திட்ட விலக்கம் அமைந்துள்ளது. திட்டவிலக்கத்தின் வாய்பாடு கணக்கியல் முறைக்கு இயைபாக அமைந்து மாறிகளின் ஏற்ற இறக்கங்களினுல் பாதிக்கப்படாததினுல் இதன் உபயோகம் மேலும் வலியுறுத்தும்படி அமைந்துவிட்டது.

ஓர் ஒழுங்கு நிகழ்வெண் பரவலாக அமைந்த மாறியின் திட்டவிலக்கம் அடியிற்காணும் வாய்பாடுகளால் சரிபார்க்கப் படுவது மிகவும் உபயோகமாக உள்ளது. இத் தன்மை உயிர்ப் புள்ளியியல் அளவுகளுக்கு மிகவும் பொருந்தும்.



படம் 3·4

$$\bar{x} \pm \sigma = 68\%$$

$$\bar{x} \pm 2\sigma = 95\%$$

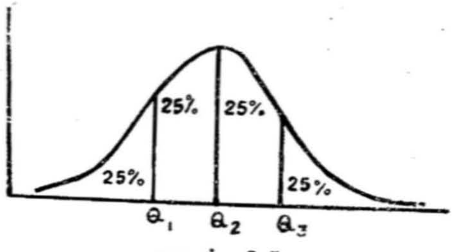
$$\bar{x} \pm 3\sigma = 99\%$$

மாறியின் கூட்டுச் சராசரிக்கு இருபுறமும் திட்டவிலக்கத் தூரம் கணக்கிடும்போது, மாறியின் மொத்த அளவுகளில் 68·27 சத அளவுகள் அமைவதாயும், இருபுறமும் இருதிட்ட விலக்கத் தூரம் 95·54 சத அளவுகளைக் கொள்வதாயும், இரு புறமும் மூன்று திட்டவிலக்கத் தூரம் 99·73 சத அளவுகளைக் கொள்வதாயும் அமைவதை உணர்வது அவசியமானதாகும். இத் தன்மை திட்டவிலக்கத்திற்கு ஒரு நிலையான பொருளைத் தருவதோடு, பெரும்பாலான இடங்களில் ஏற்படும் தவறுகளைச் சரிபார்க்கச் சாதகமாகவும் அமைந்துள்ளது.

திட்டவிலக்கத்தை மற்ற விலக்க அளவைகளைக் காட்டிலும் இலகுவாகக் கணக்கிடமுடிவதில்லை; மற்றும் எளிதாகப் புரிய வைக்க முடிவதில்லை என்பதையன்றி வேறு குறிப்பிடத்தக்க குறைபாடுகள் இந்த அளவையில் இருப்பதாகக் கூறமுடியவில்லை. மற்ற விலக்க அளவைகளைக் கணக்கிடுவதற்கான தக்க காரணங்கள் இருந்தாலன்றித் திட்டவிலக்கத்தை விலக்க அளவையாகக் கொள்வதைத் தவிர்க்கக் கூடாது.

கால்மான விலக்கம் (Quartile Deviation)

இடைநிலையின் தன்மையை ஒத்து அமைந்துள்ள வேறு இரு அளவுகள் கீழ்க் கால்மானம் (lower Quartile) அல்லது முதல் கால்மானம் (First Quartile) என்றும், மேல் கால்மானம் (Upper Quartile) அல்லது மூன்றுவது கால்மானம் (Third Quartile) என்றும் கொள்ளப்படும். இடைநிலையை (Median) இரண்டாவது கால்மானமாகக் கொண்டு சீரான அல்லது ஓரளவு சமச்சீரற்ற நிகழ்வெண் பரவலைப் பிரித்தால் நான்கு சம பாகங்கள் கிடைப்பதை உணரலாம்.



படம் 3·5

இடைநிலை, இருபுறமும் சம உறுப்புகளைக் கொண்டதாயும், முதல் கால்மானத்திற்கு இடப்புறம் 25 சத உறுப்புகளும், வலப் புறம் 75 சத உறுப்புகளும், மூன்றுவது கால்மானத்திற்குக் கீழ் 75 சத உறுப்புகளும், மேல் 75 சத உறுப்புகளும் அமையும்.

முதல் கால்மானத்திற்கும் மூன்றாவது கால்மானத்திற்கும் இடையில் 50 சத உறுப்புகள் அமைகின்றன. இவ்விரு கால் மானங்களுக்கும் உள்ள வித்தியாசத்தை $(Q_3 - Q_1)$ இடைக் கால் மான வீச்சு (inter quartile range) எனக் கூறுகிறோம். இதன் சரிபாதி அளவையே $\left(\frac{Q_3 - Q_1}{2}\right)$, கால்மான விலக்கம் (Quartile Deviation : Q. D.) அல்லது அரை இடைக் கால்மான வீச்சு (Semi-inter quartile range) எனக் கொள்ளப்படும்.

$$Q.D : \frac{Q_3 - Q_1}{2}$$

முதற் கால்மானமும், மூன்றாவது கால்மானமும் இடைநிலையைக் கணக்கிடும் யுக்தியைப் பயன்படுத்தியே கணக்கிடப் படுகின்றன. நிகழ்வெண் பரவலின் வளர் நிகழ்வெண்களை அமைத்து, முதல், மூன்றாவது கால்மானப் பிரிவு இடை வெளிகளைக் கணித்து, இவற்றின் மதிப்புகள் வரையறுக்கப்படுகின்றன. $\frac{N}{4}$-வது உறுப்பு அமையும் இடைவெளி முதல் கால்மானப் பிரிவு இடைவெளி எனவும், 3/4 N-வது உறுப்பு அமையும் இடைவெளி மூன்றாவது கால்மானப் பிரிவு இடைவெளி எனவும் கொள்ளப் படும்.

முதல் கால்மான மதிப்பிற்கான வாய்பாடு :

$$Q_1 = l + \left(\frac{\frac{N}{4} - m}{f}\right) c$$

l : முதல் கால்மானப் பிரிவு இடைவெளியின் கீழ் எல்லை மதிப்பு

m : முதல் கால்மான இடைவெளிப் பிரிவு வரை அமைந்த வளர் நிகழ்வெண்களின் மதிப்பு.

f : முதல் கால்மான இடைவெளிப் பிரிவின் நிகழ்வெண்

c : முதல் கால்மான இடைவெளிப் பிரிவின் அளவு.

மூன்றாவது கால்மான மதிப்பிற்கான வாய்பாடு

$$Q_3 : l + \left(\frac{\frac{3}{4}N - m}{f}\right) c$$

l : மூன்றாவது கால்மான பிரிவு இடைவெளியின் கீழ் எல்லை மதிப்பு.

m : மூன்ருவது கால்மான இடைவெளிப் பிரிவு வரை அமைந்த வளர் நிகழ்வெண்களின் மதிப்பு.

f : மூன்ருவது கால்மான இடைவெளிப் பிரிவின் நிகழ்வெண்.

c : மூன்ருவது கால்மான இடைவெளிப் பிரிவின் அளவு.

காட்டு :

x	f	$<$ cu fr.	உண்மை இடைவெளி	
20—21	2	2	19·5—21·5	
22—23	14	16	21·5—23·5	
24—25	229	245	23·5—25·5	
26—27	155	400	25·5—27·5 ...	Q_1
28—29	167	567	27·5—29·5	
30—31	339	906	29·5—31·5	Q_3.
32—33	76	982	31·5—33·5	
34—35	62	1044	33·5—35·5	
36—37	13	1057	35·5—37·5	
38—39	14	1071	37·5—39·5	
40—41	6	1077	39·5—41·5	
42—43	1	1078	41·5—43·5	
44—45	2	1080	43·5—45·5	
	1080			

$$\frac{N}{4} = 270 \ , \ \frac{3}{4} N = 810$$

x_{270}-வது உறுப்பு, x_{810}-வது உறுப்பு அமையும் இடைவெளிகளே முதல், மூன்ருவது கால்மான இடைவெளிப் பிரிவுகள் எனக்கொள்ளப்படும்.

முதல் கால்மானம் : Q_1

$$l : 25{\cdot}5, \quad m = 245$$

$$f = 155, \quad c = 2$$

$$Q_1 = l + \left(\frac{\frac{N}{4} - m}{f}\right) c$$

$$= 25{\cdot}5 + \left(\frac{270 - 245}{155}\right) 2$$

$$= 25{\cdot}5 + 0{\cdot}322$$

$$= 25{\cdot}822$$

மூன்றுவது கால்மானம் : Q_3

$$l = 29{\cdot}5, \quad m = 567$$

$$f : 339, \quad . \quad c = 2$$

$$Q_3 = l + \left(\frac{\frac{3}{4}N - m}{f}\right) c$$

$$= 29{\cdot}5 + \left(\frac{810 - 5679}{339}\right) 2$$

$$= 29{\cdot}5 + 1{\cdot}434 = 30{\cdot}394$$

கால்மான விலக்கம் :

$$Q.D = \frac{Q_3 - Q_1}{2}$$

$$= \frac{30{\cdot}394 - 25{\cdot}822}{2} = \frac{5{\cdot}112}{2}$$

$$= 2{\cdot}556$$

இடைநிலை (median) மதிப்பைக் காண வரைபடத்தை உபயோகப்படுத்தியது போலவே, மேலினம் அல்லது கீழின வளர் நிகழ்வெண் வளைகோடுகளை வரைந்து முதல் கால்மானம், மூன்றுவது கால்மானத்திற்கான மதிப்புகளைத் தெரிந்து கொள்ளலாம்.

y–அச்சில் $\frac{N}{4}$, $\frac{3}{4}$ N—அளவுகளுக்குச் சமமாக x–அச்சிற்கு இணைகோடுகள் வரைய வேண்டும். இவ் விணைகோடுகள் வளை கோட்டினைத் தொடுமிடத்தில், x-அச்சிற்குச் செங்குத்துக் கோடுகள் வரைய வேண்டும். இச் செங்குத்துக் கோடுகள் x-அச்சைத் தொடும் இடத்திற்கான மதிப்புகளே முதல், மூன்றுவது கால்மானங்களுக்கான மதிப்புகளாகும்.

கால்மான விலக்கம் ($Q.D$) சுலபமாகக் கணக்கிடப்படுவதோடு புரிந்து கொள்வதற்கு எளிதாயும் அமைந்துள்ளது. நிகழ்வெண் பரவலில் நிகழ்வெண் நெருக்கம் பரவலின் மத்தியில் அமையும் போதுதான் இதன் உபயோகம் வலியுறுத்தப்பட வேண்டும். சீரற்ற தன்மைத்தான இடைவெளிப் பிரிவுகள் உள்ளபோதும், திட்டவிலக்கத்தைக் கணக்கிடக் கடினமாக உள்ளபோதும், பிரிவு இடைவெளிகளின் முதல், முடிவு (First and last class intervals) எல்லைகளை அறுதியிட்டுக் கூற முடியாத சூழலிலும் இதன் உபயோகம் இன்றியமையாததாய் உள்ளது. மத்திய புள்ளியிலிருந்து இருபுறமும் மூன்று கால்மான விலக்கத் தூரம் 75 சத உறுப்புகளைப் பெற்றதாய் அமையும். சமச்சீரான அல்லது ஓரளவு சமச்சீரற்ற பரவல்களில் கால்மானமும், திட்ட விலக்கமும் $Q.D = 2/3\ s.D$ என்ற முறைப்படி அமைவதைக் காணலாம்.

சிதறல் கெழுக்கள்

மாறியின் மதிப்புகள் எந்த அலகுகளால் அளக்கப்படுகின்றனவோ, அந்த அலகுகளிலேயே சராசரி மற்றும் சிதறல் அளவைகள் குறிக்கப்படும். உயரத்தைக் குறிக்கும் சராசரி அளவை x-செ.மீ. எனவும், எடையைக் குறிக்கும் சிதறல் அளவை y கிலோ எனவும் குறிக்கப்படும். வேறுபட்ட அலகுகளைக் கொண்ட மாறிகளின் மாறுபாட்டை (Variation) ஒப்பிட, அளவைகளை மாறிகளுக்கான அலகுகளினின்றும் விடுபடச் செய்யவேண்டும். அலகுகளினின்றும் விடுபட்ட அளவைகளையே இங்குக் காண உள்ளோம்.

மாறுபாட்டுக்கெழு (Co-efficient of Variation)

ஒரு மாறியின் திட்ட விலக்கத்திற்கும், கூட்டுச் சராசரிக்கும் உள்ள விகித சமத்தைச் சதவிகிதத்தில் குறிப்பிடும் மதிப்பே மாறுபாட்டுக்கெழு எனச் சொல்லப்படுகிறது. ஒப்பிடும் தன்மைக்குப் பெரும்பாலும் உபயோகப்படுத்துவது இக் கெழுவினையே ஆகும்.

$$\text{மாறுபாட்டுக்கெழு} : \frac{\text{திட்டவிலக்கம்}}{\text{கூட்டுச்சராசரி}} \times 100$$

$$C.V = \frac{S.D}{A.M} \times 100 = \frac{\sigma x}{\bar{x}} \times 100$$

மாறுபாட்டுக் கெழுவின் பகுதியில் திட்ட விலக்கம் அமைந்துள்ளதால், திட்ட விலக்கத்தின் அளவு அதிகப்பட அதிகப்பட மாறுபாட்டுக்கெழுவின் மதிப்பும் பெரிதாக அமையும், மாறு

பாட்டுக்கெழு குறைவாக இருக்கும்போது, மாறியின் அளவுகள் கூட்டுச் சராசரியைச் சுற்றி நெருக்கமாக அமைந்திருக்கும். கூட்டுச் சராசரியின் மதிப்புப் பூச்சியத்திற்கு அருகில் அமையும் தன்மையில், இக் கெழு நம்பமுடியாத தன்மையை அடைந்து விடு வதைக் கவனத்தில் கொள்வது அவசியமாகும். இரு நிகழ்வெண் பரவல்களை ஒப்பிடும்போது குறைந்த அளவு மாறுபாட்டுக்கெழு உள்ள நிகழ்வெண் பரவல் அதிக அளவு மாறுபாட்டுக்கெழு உள்ள நிகழ்வெண் பரவலைக் காட்டிலும் நெருக்கமாகவும் சீராக வும் பரவி உள்ளதாகக் கொள்ளப்படும்.

காட்டு :

கூட்டுச் சராசரி 36·24 எனவும், திட்ட விலக்கம் 4·964 எனவும் அமைந்த முன்பு கண்ட நிகழ்வெண் பரவலுக்கான மாறு பாட்டுக் கெழுவைக் கணக்கிடுவோம்.

கொடுக்கப்பட்டவை $\bar{x} = 36{\cdot}24$, S.D $= 4{\cdot}964$

மாறுபாட்டுக்கெழு $\dfrac{\text{திட்ட விலக்கம்}}{\text{கூட்டுச் சராசரி}} \times 100$

$$\text{C.V} : \frac{\text{S.D}}{\bar{x}} \times 100$$

$$= \frac{4{\cdot}964}{36{\cdot}24} \times 100 = 13{\cdot}7$$

கோட்டம் (Skewness)

சராசரிகளின் தன்மையால் மாறிகளின் மையப் போக்கினை யும், சிதறல் அளவைகளால், சராசரியைச் சுற்றி மாறியின் அளவுகள் எவ்வாறு பரவி உள்ளள என்பதையும் அறிய முடி கிறது. நிகழ்வெண் பரவலின் குறிப்பிடத்தக்க மற்ெருரு பண்பு கோட்டமாகும். சமச்சீரான நிகழ்வெண் பரவலில், வளை கோட்டின் மையப் புள்ளியிலிருந்து இருபுறமும் ஒரே அமைப்பைக் காண முடியும். சீராக அமையாத பரவலை, கோட்ட முடைப் பரவல் எனக் கொள்கிறோம். நடைமுறையில் காணப்படும் பெரும்பாலான பரவல்கள் கோட்டமுடைப் பரவல்களாக அமை வதால், இதனை ஓர் அளவையால் குறிப்பிடும் தேவை ஏற்படு கிறது.

சமச்சீரான பரவலின் கோட்ட அளவைப் பூச்சியம் எனக் கொண்டு, கோட்ட அளவைகளின் தன்மை விளக்கப்படுகிறது. வளைகோட்டின் நீண்ட வால்புறம் வலப் புறம் நீண்டிருக்கும் தன்மையை நேர்கோட்ட முடைப் பரவல் (Positively skewed)

எனவும், நீண்ட வால்புறம் இடப்புறம் நீண்டிருக்கும் தன்மையை எதிர்கோட்ட மூடைப் பரவல் (negetively skewed) எனவும் குறிப்பிடுகிறோம்.

பியர்சனின் கோட்டக்கெழு (Karl Pearson's Co-efficiency of Skewness).

$$\text{கோட்டக்கெழு} = \frac{\text{கூட்டுச் சராசரி—முகடு}}{\text{திட்ட விலக்கம்}}$$

முகட்டினைச் சரியாகக் கணக்கிடமுடியாத இடத்தில் (கூட்டுச் சராசரி—முகடு) என்பதற்கு ஈடாக 3 (கூட்டுச் சராசரி—இடைநிலை) என்ற மதிப்பை உபயோகப்படுத்திக் கோட்டக் கெழுவினைக் கணக்கிட்டுக்கொள்ளலாம்.

$$\text{கோட்டக்கெழு} = \frac{3\ (\text{கூட்டுச் சராசரி—இடைநிலை})}{\text{திட்ட விலக்கம்}}$$

காட்டு :

இடைவெளிப் பிரிவு	நிகழ் வெண்	கீழின வளர் நிகழ் வெண்	உண்மை இடைவெளி	
25—28	8	8	24·5—28·5	
29—32	259	267	28·5—32·5	
33—36	367	634	32·5—36·5	Q_1
37—40	258	892	36·5—40·5	
41—44	72	964	40·5—44·5	
45—48	114	1078	44·5—48·5	
	1078			

இந் நிகழ்வெண் பரவலுக்கான கூட்டுச் சராசரியையும் திட்ட விலக்கமும் முன்பே கணக்கிடப்பட்டுள்ளதை அறிவோம்.

$\bar{x} : 36{\cdot}240,\ \sigma x : 4{\cdot}964.$

இடைநிலை :

$$M = l + \left(\frac{\frac{N}{2} - m}{f}\right)c$$

$$: 32{\cdot}5 + \left(\frac{539 - 267}{367}\right) 4$$

$$: 32\,5 + 2{\cdot}964$$

$$: 35{\cdot}464$$

முகடு : $Z = l + \left(\frac{f - f_1}{2f - f_1 - f_2}\right) c.$

$$= 32{\cdot}5 + \frac{(367 - 259)}{(2 \times 367 - 259 - 258)} \times 4$$

$$= 32{\cdot}5 + 1{\cdot}99$$

$$= 34{\cdot}49$$

கோட்டக்கெழு : $Sk : \frac{\text{கூட்டுச் சராசரி} - \text{முகடு}}{\text{திட்ட விலக்கம்}}$

$$: \frac{\bar{x} - z}{\sigma x}$$

$$: \frac{36{\cdot}24 - 34{\cdot}49}{4{\cdot}964} = \frac{1{\cdot}749}{4{\cdot}964}$$

$$: 0{\cdot}3523$$

கோட்டுக்கெழு :

$$: Sk : \frac{3\ (\text{கூட்டுச் சராசரி} - \text{இடைநிலை})}{\text{திட்ட விலக்கம்}}$$

$$: \frac{3\ (x - M)}{\sigma x}$$

$$: \frac{3\ (36{\cdot}24 - 35{\cdot}464)}{4{\cdot}964}$$

$$: \frac{3 \times 0{\cdot}776}{4{\cdot}964} = \frac{2{\cdot}328}{4{\cdot}964}$$

$$: 0{\cdot}469$$

கால்மானக் கோட்டக்கெழு (Quantile Co-efficient of Skewness)

சமச்சீரான பரவல்களில் முதல், இடை, மூன்றாவது கால் மானங்கள் சமதூரத்தில் அமையும் என்பதைச் சுட்டினோம்.

$$Q_2 - Q_1 = Q_3 - Q_2$$

$$Q_3 + Q_1 = 2\,Q_2$$

$$Q_3 + Q_1 - 2\,Q_2 : 0$$

நிகழ்வெண் வளைகோடு சீராக அமையாவிட்டால் $(Q_2 - Q_1) > (Q_3 - Q_2)$ அல்லது $(Q_2 - Q_1) < (Q_3 - Q_2)$ என அமையும். சீரான பரவலின் கோட்டம் பூச்சியமானதால் $Q_1 + Q_3 - 2\,Q_2$ என்பதைக் கால்மானக் கோட்ட அளவையாகக் கொள்ளலாம்.

கால்மானக் கோட்டக்கெழு : $\beta_1 = \dfrac{Q_3 + Q_1\ 2\,Q_2}{Q_3 - Q_1}$

$\beta_1 = 0$ கோட்டமற்ற பரவல்

$\beta_1 > 0$ நேர்கோட்டமுடைப் பரவல்

$\beta_1 < 0$ எதிர் கோட்டமுடைப் பரவல்

காட்டு :

$Q_1 = 25{\cdot}822,\ Q_2 = 29{\cdot}177,\ Q_3 = 30{\cdot}934$ என அமையும் நிகழ்வெண் பரவலுக்கான கோட்டக்கெழுவைக் கணக்கிடுவோம்.

$$\beta_1 = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

$$= \frac{30{\cdot}934 + 25{\cdot}822 - 2\,(29{\cdot}177)}{30{\cdot}934 - 25{\cdot}822}$$

$$= \frac{-\ 1{\cdot}598}{5{\cdot}112} = -\ 0{\cdot}3126.$$

கோட்டக் கெழுக்களின் மதிப்பிற்கான எல்லைகளைக் குறிப்பிடுவது கடினமானதாயினும், ஓரளவு சமச்சீரற்ற பரவல்களுக்கான கோட்டக்கெழுக்கள் மிக அதிகமான மதிப்பைப் பெறுத தன்மைத்தாய் உள்ளது. பெரும்பாலான இடங்களில் பியர்சனின் கோட்டக்கெழு — 3-லிருந்து + 3 வரையும், கால்மானக் கோட்டக்கெழு — 1-லிருந்து + வரையும் அமைவதைக் காணலாம். கோட்டக் கெழுவின் மதிப்பிற்கேற்பக் கோட்டத்தன்மை விளக்கப்படும்.

காட்டு :

அடியிற்காணும் நிகழ்வெண் பரவலுக்கான பியர்சனின் கோட்டக்கெழு, கால்மானக்கோட்டக் கெழுக்களைக் கணக்கிடுவோம்.

இடைவெளிப் பிரிவு	நிகழ் வெண்	பிரி வெண்	u_i	$u_i f_i$	$u_i^2 f_i$	வளர் நிகழ்வெண்
35—39	14	37	—3	—42	126	14
40—44	60	42	—2	—120	240	74
45—49	105	47	—1	—105	105	179
50—54	138	52	0	—267	·	317
55—59	129	57	1	129	129	446
60—64	85	62	2	170	340	531
65—69	28	67	3	84	252	559
70—74	15	72		64	240	574
				443		
	574			+176	961	

கூட்டுச்சராசரி : $\bar{x} = 52 + 5\left(\frac{176}{574}\right)$

$= 52 + 5\,(0{\cdot}307)$

$= 53{\cdot}535$

இடைநிலை : M $= 49{\cdot}5 + \left(\frac{287 - 179}{138}\right)5$

$= 49{\cdot}5 + 3{\cdot}913$

$Q_2 = 53{\cdot}413$

முகடு : Z $= 49{\cdot}5 + \left(\frac{138 - 105}{276 - 105 - 129}\right)5$

$= 49{\cdot}50 + 3{\cdot}93$

$= 53{\cdot}43$

திட்டவிலக்கம் : $5\sqrt{\frac{961}{574} - (\cdot 307)^2}$

$$= 5 \times 1{\cdot}549$$
$$= 7{\cdot}745$$

$$\text{பியர்சனின் கோட்டக்கெழு} = \frac{\text{கூட்டுச்சராசரி} - \text{முகடு}}{\text{திட்டவிலக்கம்}}$$

$$= \frac{53{\cdot}535 = 53{\cdot}4}{7{\cdot}745}$$

$$Sk = 0{\cdot}01$$

$$\text{பியர்சனின் கோட்டக்கெழு} = \frac{3(\text{கூட்டுச்சராசரி} - \text{இடைநிலை})}{\text{திட்டவிலக்கம்}}$$

$$= \frac{3\ (53{\cdot}535 - 53{\cdot}413)}{7{\cdot}745}$$

$$Sk = 0{\cdot}047$$

$$Q_1 = 44{\cdot}5 + \left(\frac{143{\cdot}5 - 74}{105}\right)5 \qquad Q_3 = 54{\cdot}5 + \left(\frac{430{\cdot}5 - 317}{129}\right)5$$

$$= 44{\cdot}5 + 3{\cdot}3095 \qquad\qquad = 54{\cdot}5 + 4{\cdot}4$$

$$= 47{\cdot}81 \qquad\qquad = 58{\cdot}9$$

$$\text{கால்மானக் கோட்டக்கெழு} : \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

$$= \frac{58{\cdot}90 + 47{\cdot}81 - 2\ (53{\cdot}41)}{58{\cdot}9 - 47{\cdot}81}$$

$$= -\ {\cdot}01$$

பயிற்சி

விளைபலன்	பாத்திகள்
2·7—2·9	4
2·9—3·1	15
3·1—3·3	20
3·3—3·5	47
3·5—3·7	63
3·7—3·9	78
3·9—4·1	88
4·1—4·3	69

1. 1/500 ஏக்கர்ப் பரப்புள்ள பாத்திகளின் விளைபலனைத் தரும் மேற்கண்ட நிகழ்வெண்பரவலுக்கான இடைநிலை, முதல், மூன்றாவது கால்மான அளவுகளைக் காண். வரைபடம் மூலம் கிடைக்கும் அளவினுக்குள்ள வித்தியாசத்தைக் குறிப்பிடு

2. 12 பகடைகளை 4096 தடவை வீசி 4,5,6 கிடைப்பதை வெற்றிகளின் எண்ணிக்கையாகக் கொண்ட முடிவுகள் தரப் பட்டுள்ளன. கூட்டுச்சராசரி, திட்டவிலக்கம் ஆகியவற்றைக் கணக்கிடு.

வெற்றி	0	1	2	3	4	5	6
நிகழ்வெண்	—	7	60	1	430	731	948
வெற்றி	7	8	9	10	11	12	மொத்தம்
நிகழ்வெண்	847	536	257	71	11	—	4096

3. x, y எனும் இருமாறிகளுக்கான மதிப்புகள் தரப் பட்டுள்ளன. இவ் விருமாறிகளுக்கான மாறுபாட்டுக் கெழுக்களைக் கணக்கிட்டு, அவற்றின் வேறுபாட்டின் தன்மையை விளக்கு.

x	89	86	74	65	63	66	77	65	72	79
y	82	91	84	75	73	72	70	75	78	84

4. கொடுக்கப்பட்டுள்ள பரவலுக்கான திட்டவிலக்கத்தைக் கணக்கிடு.

இடைவெளிப் பிரிவு மேல் எல்லை	கீழின வளர் நிகழ்வெண்
10	25
20	40
30	60
40	75
50	95
60	125
70	190
80	240

5. அடியிற்காணும் நிகழ்வெண்பரவலுக்கு இடைநிலை, கீழ்க்கால்மானம், மேல்கால்மான அளவுகளைக் கணக்கிடு.

மதிப்பெண்	மாணவர்கள்
0—4	10
4—8	12
8—12	18
12—14	7
14—18	5
18—20	8
20—25	4
25க்கு மேல்	6

6. 100 மாணவர்களின் அறிவுக் குறியீட்டெண்களுக்கான (Intelligence auotient) நிகழ்வெண்பரவல் கொடுக்கப்பட்டுள்ளது. இதற்கான இடைநிலை மதிப்பைக் கணக்கிடு.

x :	55—64	65—74	75—84	85—94	95—104
f :	1	2	9	22	33

x :	105—114	115—124	125—134	135—144
f :	22	8	2	1

7. நூறு குழந்தைகளின் உயரம் (அங்) கொடுக்கப்பட்டுள்ளன. மாறுபாட்டுக் கெழுவைக் கணக்கிடு.

உயரம்	60	61	62	63	64	65	66	67	68
நிகழ்வெண்	2	0	15	29	25	12	10	4	3

8. 25 மாணவர்கள் உயிர்ப் புள்ளியியல், உயிரியல் ஆகிய இரு தேர்வுகளில் பெற்ற மதிப்பெண்கள் கொடுக்கப்பட்டுள்ளன. இதற்கான இடைநிலை மதிப்பைக் கணக்கிட்டு, எந்தப் பாடத்தில் மாணவர்தம் திறன் அதிகமாக உள்ளது என்பதைக் காரணத்துடன் விளக்கு.

மாணவர்	1	2	3	4	5		7	8	9
உயிர்ப் புள்ளியியல்	29	65	33	45	51	72	48	33	42
உயிரியல்	36	30	38	39	64	50	46	15	42
மாணவர்	10	11	12	13	14	15	16	17	
உயிர்ப் புள்ளியியல்	25	28	35	46	47	60	30	32	
உயிரியல்	10	72	33	80	44	85	20	32	
மாணவர்	18	19	20	21	22	23	24	25	
உயிர்ப் புள்ளியியல்	52	54	56	58	49	38	40	46	
உயிரியல்	25	55	28	53	35	40	62	58	

9. உயிர்ப் புள்ளியியல் தேர்வில் பெற்ற மதிப்பெண்கள் நிகழ்வெண்பரவலாகக் கொடுக்கப்பட்டுள்ளன. இதற்கான கூட்டுச்சராசரி, சிதறல் கெழுவினைக் கணக்கிடு.

மதிப்பெண்	மாணவர்கள்	மதிப்பெண்	மாணவர்கள்
10—20	1	60—70	12
20—30	2	70—80	16
30—40	3	80—90	10
40—50	5	90—100	4
50—60	7	மொத்தம்	60

10. 1941 மக்கட் கணக்குப்படி பரோடா மாநிலத்தின் மணமான பெண்களின் நிகழ்வெண்பரவல் தரப்பட்டுள்ளது. இதற்கான இடைநிலை, கால்மான விலக்கத்தைக் கணக்கிடு.

இடைவெளிப் பிரிவு	நிகழ்வெண்	இடைவெளிப் பிரிவு	நிகழ்வெண்
0—5	3	40—45	963
5—10	31	45—50	702
10—15	410	50—55	531
15—20	1809	55—60	317
20—25	2446	60—65	156
25—30	2223	65—70	59
30—35	1723	70—75	37
35—40	1292	75—80	38

11. எலும்பு வளர்ச்சிக்காக, ஒவ்வொரு ஜோடிப் பன்றிக்கும் சுண்ணும்புச்சத்து, எலும்புச்சத்துக் கலந்த இருவகை உணவுகள் கொடுக்கப்பட்டு அடியிற்காணும் முடிவுகள் குறிக்கப்பட்டுள்ளன. மாறுபாட்டுக்கெழுவினைக் (Co-efficient of variation) கணக்கிட்டு உணவுகளின் தன்மையை விவரி.

ஜோடி எண்	சுண்ணும்புச்சத்து	எலும்புச்சத்து
1	49·2	51 5
2	53·3	54·9
3	50·6	52·2
4	52·0	53·3
5	46·8	51·6
6	50·5	54·1
7	52·1	54·2
8	53·0	53·3

12. பின்வரும் நிகழ்வெண்பரவலின் இடைநிலை, கால்மான விலக்கத்தைக் கணக்கிடு.

இடை வெளிப் பிரிவு	30—39	40—49	50—59	60—69	70—79	80—89	90—99	மொத்தம்
நிகழ் வெண்	2	3	11	20	32	25	7	100

13. அடியிற்காணும் நிகழ்வெண்பரவலுக்கான கால்மான விலக்கத்தைக் கணக்கிடு.

x_i	f_i
55—64	3
65—74	21
75—84	78
85—94	182
95—104	305
105—114	209
115—124	81
125—134	21
135—140	5

14. ஒரு தேர்வில் 1000 மாணவர்களின் உயிர்ப் புள்ளியியலில், புள்ளியியலில் பெற்ற மதிப்பெண்கள் தரப்பட்டுள்ளன.

மதிப்பெண்கள்	0—9	10—19	20—29	30—39	40—49	50—59
உயிர்ப் புள்ளியியல்	1	3	19	131	146	340
புள்ளியியல்	20	41	98	149	192	197

மதிப்பெண்கள்	60—69	70—79	80—89	90—99	மொத்தம்
உயிர்ப் புள்ளியியல்	134	20	4	2	1000
புள்ளியியல்	152	97	94	10	1000

உயிர்ப் புள்ளியியல், புள்ளியியல் தேர்வுகளுக்கான மாறுபாட்டுக் கெழுக்களைக் கணக்கிடு.

4. நிகழ்திறன்

(Probability)

நிகழ்ச்சிகள் நிகழ்வதைப் பொது அறிவுக்கு முரண்படாமல் அளவிட்டுக் கூறுவதே நிகழ்திறனுகும். புள்ளியியல் தன்மைப்படி நிகழ்திறனுனது கொடுக்கப்பட்ட முழுமைத் தொகுதியின் ஒரு குறிப்பிட்ட செயல் நிகழ்வதற்கான வாய்ப்பினைக் குறிக்கும். இது ஒரு பண்பாக இருக்கலாம், தனிப்பட்டவர்களின் பாலினைக் குறிக்கலாம். பேருந்து வண்டியின் ஒரு பாகத்தையோ அல்லது செய்து முடிக்கப்பட்ட பேருந்து வண்டியின் பொருட்கூட்டிலுள்ள பழுதுபட்ட பொருள்களையோ குறிக்கலாம். மனிதனின் ஓர் உறுப்பின் பண்பைக் குறிக்கலாம் அல்லது ஒரு குடும்பத்திலுள்ள நபர்களின் எண்ணிக்கையையோ அல்லது அதன் வருமானத்தின் ஓர் இடைவெளியையோ குறிக்கலாம்.

நிகழ்திறனைக் குறிப்பிட்ட சில வார்த்தைகளால் அறுதியிட்டுக் கூறுவது கடினம். அதன் தன்மையை அனுபவத்தாலும், நிகழ்ச்சிகள் நிகழ்வனவற்றைக் கவனிப்பதாலும், அகக் காட்சியால் உணர்வதாலும் சரிபார்க்க முடியும். அதன் அடிப்படைக் கருத்துகள் பொது அறிவின் சக்தியால் அமைக்கப்பட்டவை. ‘இன்று மாலை மழை பெய்யலாம்,’ ‘நாளை விடுமுறையாக இருக்கலாம்’ என்ற கருத்துகள் சாதாரணமாகப் பரிமாறப்படுவதைக் காண்கிறோம். இதன்பால் நமக்கேற்படும் நம்பிக்கை சொல்பவர் தம் நோக்கத்தாலும், அவர்தம் தன்மையாலும் அமையப் பெறுகிறது. இப் பேச்சு வழக்கில் பெறப்படும் கருத்துகள் முடிவான தன்மை பெறுதனவாகவோ, ஊகத்திற்கு இடம் வைப்பனவாகவோ அமைகின்றன. முடிவான தன்மையும், ஊக அளவையும் நாம் சொல்லும் கருத்துகட்கோ அல்லது நம்பும் கருத்துகட்கோ நாம் கொடுக்கும் நம்பிக்கையின் அளவைக் குறிக்கும். ஊகத்திற்கு இடம் கொடுக்கும் தன்மைத்தான இடங்களில் நிகழ்திறனை உபயோகப்படுத்துகிறோம். எண்களால் குறிக்க முடியும், சந்தேகத்

திற்கிடமான சூழலிலேயே நிகழ்திறன் பயன்படுகிறது. நிகழ் திறனை அனுமானிக்க முடியும் இடங்களில் வாயால் சொல்லப் படும் ஊக அளவைகளால் நின்றுவிடுவது, அறிவியலுக்கு ஏற் காத ஒன்றுவதை நன்கறிவோம். எண்களால் குறிக்கப்படும் யூக அளவை, செய்திகளை நம்பும் தன்மை, ஒப்பிடும் முறை, உணரப் படும் வழி ஆகியவற்றைச் செம்மைப்படுத்தவோ அல்லது வலுப் படுத்தவோ செய்கிறது.

ஒரு பகடையை உருட்டும்போது அதன் ஆறு பக்கங்களில் ஏதேனும் ஒன்று மேற்புறமாக விழும். பழுதில்லாத ஒருபுறச் சார்பற்ற சிறந்த பகடையில் எந்தப் பக்கம் வரும் என்பதை நிகழ்ச்சி நிகழ்வதற்கு முன்பே கூற முடியாது. பகடை உருட்டு வதை ஒரு நிகழ்ச்சிக்கு ஒப்பிட்டு, நிகழ்ச்சிகளுக்கான சில அடிப் படைக் குணங்களைப் பார்ப்போம்.

நிகழ்ச்சிகள் தனித்தனியே நிகழ்வதற்கான வாய்ப்பு, சமமாக இருக்கும் நிகழ்ச்சிகளைச் சமவாய்ப்புடை (equally likely) நிகழ்ச்சி கள் எனக் கொள்கிறோம். சமவாய்ப்புடைய நிகழ்ச்சிகளில் ஒரு நிகழ்ச்சியைக் காட்டிலும் இன்னொரு நிகழ்ச்சி நிகழ்வதற்கான காரணம் நம்மால் ஊகிக்க முடியாதபடி இருக்கும். நன்றாகக் கலக்கப்பட்ட ஒரு சீட்டுக் கட்டிலிருந்து எடுக்கப்படும் ஒரு சீட்டு அதிலுள்ள 52 சீட்டுகளில் ஏதேனும் ஒன்றாக இருக்கும். சீட்டு எடுப்பதை ஒரு நிகழ்ச்சியாகக் கொண்டால் 52 சீட்டுகளும் சம வாய்ப்பைப் பெற்று, ஒரு சீட்டைக் காட்டிலும் மற்றொரு சீட்டு வருவதற்கான காரணமின்றிச் சமவாய்ப்புடை நிகழ்ச்சிகளாக அமைந்துள்ளதைக் காணலாம்.

இரு நிகழ்ச்சிகளில் ஏதாவதொன்று நிகழும்போது மற்றொரு நிகழ்ச்சி நிகழ முடியாத தன்மைத்தாயின், அவ்விரு நிகழ்ச்சிகளும் ஒன்றையொன்று விலக்கும் தன்மைத்தான (mutually exclusive) நிகழ்ச்சிகள் எனக் கொள்கிறோம். பகடையை உருட்டும்போது ஆறு எண்களில் (பக்கங்களில்) ஏதேனும் ஒன்று விழுவது மற்ற எண்கள் வருவதைத் தடுப்பதால் அவை ஒன்றையொன்று விலக்கும் தன்மைத்தான நிகழ்ச்சிகளாக அமைகின்றன.

நிகழ்ச்சிகள் நிகழும் எல்லா வழிகளையும் ஆராய்ந்து அதன் முழுமைத் தொகையையும் சேர்க்கும் நிகழ்ச்சிகள் பூரண தன்மைத்தான நிகழ்ச்சிகள் (exhaustive events) எனப்படும்.

ஒரு நிகழ்ச்சி நிகழ்வதை உறுதிப்படுத்தும் நிகழ்ச்சிகளையே, அந் நிகழ்ச்சியின் சாதக நிகழ்ச்சிகள் (favarable events) எனக்

கொள்கிறோம். புதிதாகப் பிறந்த 10 குழந்தைகளில் 4 பெண்ணு கவும் 6 ஆணுகவும் இருந்தால், ஆண் குழந்தை பிறந்ததற்கான சாதக நிகழ்ச்சிகள் 6 எனக் கொள்கிறோம்.

ஒரு நிகழ்ச்சி ஏற்படுவது மற்றொரு நிகழ்ச்சி ஏற்படுவதைப் பாதிக்காமலிருக்குமாளுல், அந் நிகழ்ச்சிகள் சார்பிலா நிகழ்ச்சிகள் **(independent events)** எனப்படும்.

நிகழ்திறனை (probability) இரு வழிகளில் கணக்கிடும் முறை பழக்கத்தில் உள்ளது. (a) கணக்கியல் நிகழ்திறன் (நிகழ்ச்சி நிகழ்வதற்கு முன்னமே அனுமானிக்கப்படும் நிகழ்திறன்); (b) புள்ளியியல் நிகழ்திறன் (நிகழ்ச்சி நிகழ்ந்தபின் கணக்கிடப் படும் நிகழ்திறன்).

கணக்கியல் நிகழ்திறன் (Mathematical Probability)

பூரணத்தன்மைத்தான, சமவாய்ப்புடைய, ஒன்றையொன்று விலக்கும் 'n' நிகழ்ச்சிகளில், A என்ற நிகழ்ச்சிக்கு m முறைகள் சாதகமாக இருக்குமாயின், A என்ற நிகழ்ச்சி நிகழ்வதற்கான வாய்ப்பினை p (A) எனக் குறிப்பிட்டு m/n என்ற பின்னத்தால் அளவிடுகிறோம்.

A-என்ற நிகழ்ச்சி நிகழ்வதற்கான நிகழ்திறன்

$$p(A) = \frac{\text{A என்ற நிகழ்ச்சிக்குச் சாதகமான நிகழ்ச்சிகள்}}{\text{மொத்த நிகழ்ச்சிகள்}}$$

$$= \frac{m}{n} = p$$

சாதகமான நிகழ்ச்சிகளை ஒரு நிகழ்ச்சி நிகழ்வதற்கான வெற்றிகளாகவும், மற்றவற்றைத் தோல்விகளாகவும் கருதுவது மரபு. p-ன் மதிப்பு ஒன்றுக்கு மேற்படாமல் நேரெண்ணுக அமைவதைக் காணமுடியும். m வழிகளில் வெற்றிபெறும் A எனும் நிகழ்ச்சி $(n-m)$ வழிகளில் தோல்வியடைகிறது. தோல்விக் கான நிகழ்திறனை 'q' எனக் குறிப்பிடுவோமாயின்,

$$q = \frac{n-m}{n} = 1 - \frac{m}{n} = 1 - p$$

$$q = 1 - p$$

$$p + q = 1$$

இங்கே வெற்றி, தோல்வி (வெற்றியல்லாதது) என்ற சொற்கள் பொதுநிலையில் உபயோகப்படுத்தப்படுகின்றன. இதனையே ஒரு குறிப்பிட்ட பண்பு அமையப்பெற்றவற்றை வெற்றிகளாகவும், மற்றவற்றைத் தோல்விகளாகவும் கொள்ளலாம். ஒரு கிலோ விதையில் பழுதானவற்றை வெற்றிகளாகவும், நல்ல விதைகளைத் தோல்விகளாகவும் கணக்கிடலாம். இவ்விடத்தில் ஒரு கிலோ விதையில் உள்ள பழுதான விதைகளையே நாம் கணக்கிட எண்ணுவதால், பழுதானவற்றை (பழுதானது என்ற பண்பினைப் பெற்றவற்றை) வெற்றிகளாகக் கொள்கிறோம்.

காட்டு :

கொடுக்கப்பட்டுள்ள பிறப்பு இறப்புப் பதிலிருந்து ஆண் மகவு பிறப்பதற்கான நிகழ்திறனைக் கணக்கிடுவோம்.

ஓர் ஆண்டில் பதிவான மொத்த ஆண் குழந்தைகள்: 7,05,593.
அதே ஆண்டில் பதிவான மொத்தக் குழந்தைகள் : 13,73,438.

குழந்தை பிறப்பதை ஒரு நிகழ்ச்சியாகக் (event) கொள்வோம். ஆண் குழந்தைகளைச் சாதக நிகழ்ச்சிகளாகவும், (favarable events) மொத்தக் குழந்தைகளை மொத்த நிகழ்ச்சிகளாகவும் (n) கொண்டு நிகழ்திறனைக் கணக்கிடலாம்.

$$p_{M}(A) = \frac{m}{n} = \frac{705593}{1373438} = 0{\cdot}5137$$

ஆண் மகவு பிறப்பதற்கான நிகழ்திறன் = 0·5137

காட்டு :

மூன்று கோடி மணத் தம்பதிகள் கூட்டொன்றில் ஒவ்வொரு சோடியிலிருந்தும் ஒருவரை ராண்டமாக எடுத்தால், எடுக்கப் பட்ட மூவருமே ஒரே பாலினைச் (sex) சேர்ந்தவர்களாக இருப்பதற்கான நிகழ்திறனைக் காண்போம்.

மூன்று சோடி மணத் தம்பதிகளையும்

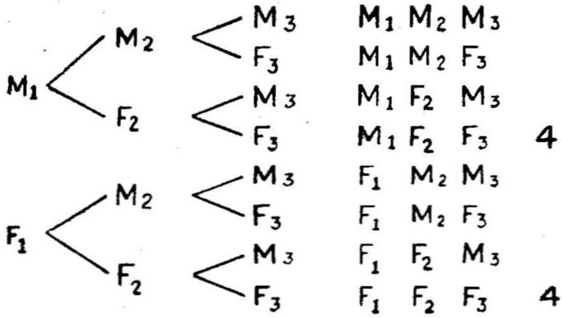
$$\begin{matrix} M_1 & M_2 & M_3 \\ F_1 & F_2 & F_3 \end{matrix}$$

எனக்கொள்வோம்.

எடுக்கப்பட்ட மூவரும் ஒரே பாலினைச் சேர்ந்தவர்களாக இருக்க மூவரும் ஆணாகவோ அல்லது பெண்ணாகவோ இருக்க வேண்டும். இதற்கான நிகழ்ச்சிகளே நிகழ்ச்சி நிகழ்வதற்கான சாதக நிகழ்ச்சிகளாகும்.

சாதக நிகழ்ச்சிகளாக $(M_1 \; M_2 \; M_3)$ மற்றும் $(F_1 \; F_2 \; F_3)$ என்ற இரண்டு நிகழ்ச்சிகள் கிடைக்கின்றன.

நிகழ்ச்சிகள் நிகழ்வதற்கான மொத்த வாய்ப்புகள் :



$2 \cdot 2 \cdot 2 = 2^3 = 8$ என அமைகின்றன.

$$\left.\begin{array}{l}\text{எடுக்கப்பட்ட மூவரும் ஒரே பாலினைச் சேர்ந்த} \\ \text{வர்களாக இருப்பதற்கான நிகழ்திறன்}\end{array}\right\} = \frac{2}{8} = \frac{1}{4}$$

கணக்கியல் அல்லது நிகழ்ச்சி நிகழ்வதற்குமுன் கணக்கிடப் படும் நிகழ்திற நியதியில் சில தவிர்க்கமுடியாத குறைபாடுகள் உள. நிகழ்ச்சிகள் நிகழ்வதற்கான மொத்த வாய்ப்புகள் எண்ணிலி யாயின் (infinite) கணக்கியல் நிகழ்திற அளவை மாற்றம் செய்யும் நிலைமைத்தாய் முடிகிறது. நேரெண்களின் கூட்டிலிருந்து இரட்டைப்படை எண்களை எடுப்பதற்கான நிகழ்திறன் $\frac{1}{2}$. இம் முடிவு போலியாக உள்ளதை விளக்குவோம். முதல் 20 அல்லது 200 எண்கள் நிகழ்திற மதிப்பை $\frac{1}{2}$ என்ற நிலையைச் செய்யும். எண்கள் இருக்கும் முறை 1, 3, 2; 5, 7, 6; 9, 11, 10.,...என்ற தன்மைத்தாயின், முதல்சோடி ஒற்றைப்படை பின்பு இரட்டைப் படை எண் அடுத்த ஒரு சோடி ஒற்றைப்படை பின்பு ஓர் இரட்டைப்படை எண், அடுத்த ஒரு சோடி ஒற்றைப்படை எண்களை அடுத்து ஓர் இரட்டைப்படை எண் .....என இம்முறை தொடர்ந்தால், நேரெண்களின் கூட்டிலிருந்து இரட்டைப்படை எண் எடுக்கும் வாய்ப்பு $\frac{1}{3}$ என அமைகிறது. எண்களை அடுக்கும் முறை, $\frac{m}{n} = p$ ஐ ஒரு முடிவான உருவைப் பெற்று நிகழ்திற அளவைக் குறிப்பிடும் முறையில் நிலையற்ற தன்மையை உண் டாக்குகிறது.

நிகழ்ச்சிகள் சமவாய்ப்பினை உடைத்தாக அமையாவிட்டால், கணக்கியல்முறை நிகழ்திறன் இன்னும் சிக்கலான நிலையை அடைகிறது. ஒருசார்புடைய (தலைப்பக்கம் கனமான) நாணயத்தைச் சுண்டுவதால் தலை விழுவதற்குச் சாதகமான நிகழ்ச்சிகள் அதிகப்படுவதை உணர முடிகிறது. இதுபோன்ற நிகழ்ச்சிகளே நடைமுறையில் அதிகம் காணப்படுவதைத் தெரிய வருகிறோம். இவ்விடங்களில் அமையும் நிகழ்திற அளவையை நம்மால் கணக்கிட முடியாவிட்டாலும், அதற்கென்று ஓர் அளவு இருப்பதை மறுக்க முடியாது. இந் நிலையில் கணக்கியல் நிகழ்திறன், வாய்ப்பை அளந்துகூறும் நிலையில் நிலையற்ற தன்மையை அடைகிறது. நிகழ்ச்சிகள் நிகழ்வதற்கு முன்பே இதனைக் கணக்கிட முடியாத நிலையில் உள்ளோம். நிகழ்ச்சிகள் நிகழ்ந்தபின் முடிவுக்கு வரும் புள்ளியியல் நிகழ்திறன் இதற்குத் துணையாக அமைகிறது.

புள்ளியியல் நிகழ்திறன் (Statistical or Empirical definition of probability)

நிலையான சூழலில் ஒரு நிகழ்ச்சியைத் திரும்பத் திரும்பச் செய்வதால் கிடைக்கும் சாதக நிகழ்ச்சிகளுக்கும் மொத்த நிகழ்ச்சிகளுக்குமான பின்னம் புள்ளியியல் நிகழ்திறனாகும். மொத்த நிகழ்ச்சிகளை அதிகப்படுத்த அதிகப்படுத்த இது ஒரு நிலையான எண்ணாக அமையும். இங்கு நிகழ்திறனானது நிகழ்ச்சிகள் நிகழ்ந்தபின் கணக்கிடப்படுகிறது.

ஒரு தொடர்ச்சியான n நிகழ்ச்சிகளில் A என்ற நிகழ்ச்சி m வழிகளில் நிகழ்ந்திருந்தால், இந் நிகழ்ச்சிக்கான புள்ளியியல் நிகழ்திறன் (Statistical probability)

$$P(A) = p = \lim_{n \to \infty} \left(\frac{m}{n}\right)$$

கணக்கியல்முறையைப் பெற்றறி முறையாகவும் (method of deduction), புள்ளியியல் முறையைப் பட்டறி முறையாகவும் (method of induction) கொள்ளலாம். நிகழ்ச்சிகள் நிகழ்வதை முன்னமே தீர்மானிக்க முடியாத நிலையில், சோதனை மூலம் உண்மையாக ஏற்படும் முடிவுகளைக்கொண்டு புள்ளியியல் நிகழ்திறன் கணக்கிடப்படுகிறது. புள்ளியியல் நிகழ்திறனின் இறுதி மதிப்பே கணக்கியல் நிகழ்திற மதிப்பின் லட்சிய எண்ணாகும்.

A_1, A_2, A_3 ...... A_n என்பனவற்றை 'n' நிகழ்ச்சிகளாகக் கொண்டு, சில குறியீடுகளை அமைப்பது புள்ளியியல் நியதியை விளக்குவதற்குச் சாதகமாக உள்ளது.

$P(A_i)$ — A_i என்ற நிகழ்ச்சி நிகழ்வதற்கான வாய்ப்பினைக் குறிக்கும். $P(\overline{A_i})$ — A_i என்ற நிகழ்ச்சி நிகழாமைக்குரிய வாய்ப்பைக் குறிக்கும். $\overline{A_i}$ என்ற நிகழ்ச்சி A_i என்ற நிகழ்ச்சிக்கு எதிரானதாக அமைகிறது. இவ்விரு நிகழ்ச்சிகளும் பூரணத்துவம் கொண்ட ஒன்றை ஒன்று விலக்கும் தன்மைத்தாய் அமைந்து

$P(A + \overline{A_i}) = 1$ என அமையும்.

$P(A_i + A_j)$ — A_i, A_j ஆகிய இரண்டு நிகழ்ச்சிகளில் ஏதேனும் ஒரு நிகழ்ச்சி நிகழ்வதற்கான நிகழ்திறனைக் குறிக்கும்.

$P(A_1 + A_2 \ldots + A_n)$ — $A_1\ A_2 \ldots\ A_n$ ஆகிய n நிகழ்ச்கல்சிளி ஏதேனும் ஒன்று நிகழ்வதற்கான நிகழ்திறனைக் குறிக்கும்.

$P(A_i\ A_j)$ ...... A_i, A_j ஆகிய இரு நிகழ்ச்சிகளும் சேர்ந்து நிகழ்வதற்கான நிகழ்திறனைக் குறிக்கும்.

$P(A_i/A_j)$ ...... A_j நிகழ்ந்தபின் A_i நிகழ்வதற்கான நிகழ்திறனைக் குறிக்கும்.

$P(A_i/A_j\ A_k)$ ...... A_j, A_k ஆகிய இரு நிகழ்ச்சிகளும் நிகழ்ந்தபின் A_i நிகழ்வதற்கான நிகழ்திறனைக் குறிக்கும்

$P(A_i/A_j)$ மற்றும் $P(A_i/A_j\ A_k)$, நிபந்தனை நிகழ்திறன்களாகும் (conditional probabilities). ஒரு நிகழ்ச்சி நிகழ்ந்தபின் மற்றொரு நிகழ்ச்சி நிகழ்வதற்கான நிகழ்திறனைக் கணக்கிடுவது நிபந்தனை நிகழ்திறனாகும். $A_j\ A_k$ ஆகிய நிகழ்ச்சிகளின் நிகழ்வைத் தெரிந்து A_i-ன் நிகழ்திறன் கணக்கிடுவதை $P(A_i/A_j\ A_k)$ குறிக்கும்.

நிகழ்திற நியதிகள்

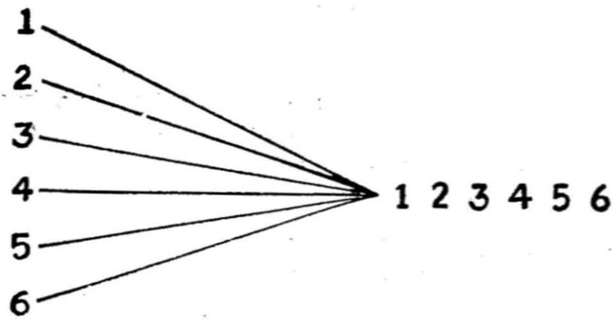
நியதி 1 : ஒன்றையொன்று விலக்கும் தன்மைத்தான A_1, A_2 என்ற இரு நிகழ்ச்சிகளில் ஏதாவதொன்று நிகழ்வதற்கான நிகழ் திறன், நிகழ்ச்சிகள் ஒவ்வொன்றும் தனித்தனியே நிகழ்வதற்கான நிகழ்திறன்களின் கூட்டுத் தொகைக்குச் சமமாகும்.

$$P(A_1 + A_2) = P(A_1) + P(A_2).$$

காட்டு :

இரு பகடைகளை உருட்டும்போது ஏழு அல்லது எட்டுக் கிடைப்பதற்கான நிகழ்திறனைக் கணக்கிடுவோம்.

இரு பகடைகளை உருட்டுவதால் ஏற்படும் மொத்த நிகழ்ச்சிகள் 36 ஆகும். முதல் பகடையில் விழும் ஒன்று, இரண்டாவது பகடையின் ஏதாவதோர் (1,2,3,4,5,6) எண்ணுடனும் .....முதல் பகடையில் விழும் ஆறு, இரண்டாவது பகடையின் ஏதாவதோர் எண்ணுடனும் சேர்ந்து மொத்த நிகழ்ச்சிகள் 36 என அமைகின்றன.



ஏழு கிடைப்பதற்கான சாதக நிகழ்ச்சிகள், முதல் பகடையின் ஒன்றும் இரண்டாவது பகடையின் ஆறும், முதல் பகடையின் இரண்டும் இரண்டாவது பகடையின் ஐந்தும் ..... முதல் பகடையின் ஆறும் இரண்டாவது பகடையின் ஒன்றுமாக ஆறு வழிகளில் நிகழ்கின்றன,

முதல் பகடை	இரண்டாவது பகடை
1	6
2	5
3	4
4	3
5	2
6	1

எட்டுக் கிடைப்பதற்கான சாதக நிகழ்ச்சிகள் முதல் பகடையின் இரண்டும் இரண்டாவது பகடையின் ஆறும் ...... முதல் பகடையின் ஆறும் இரண்டாவது பகடையின் இரண்டுமாக ஐந்து உள.

முதல் பகடை	:	2	3	4	5	6
இரண்டாவது பகடை	:	6	5	4	3	2

ஏழு அல்லது 8 கிடைப்பதற்கான நிகழ்திறன் :

$$p = \frac{\text{சாதக நிகழ்ச்சிகள்}}{\text{மொத்த நிகழ்ச்சிகள்}}$$

$$P(A_1 + A_2) = P(A_1) + P(A_2)$$

$$= \frac{6}{36} + \frac{5}{36} = \frac{11}{36}$$

நியதி 2: A_1, A_2 என்ற இரு சார்பிலா நிகழ்ச்சிகள் சேர்ந்தாற்போல் நிகழ்வதற்கான நிகழ்திறன் இவ்விரு நிகழ்ச்சிகளும் தனித்தனியே நிகழ்வதற்கான நிகழ்திறன்களின் பெருக்குத் தொகைக்குச் சமமாகும்.

$$P(A_1 \cdot A_2) = P(A_1) \cdot P(A_2)$$

காட்டு :

பை ஒன்றுக்கு 6 சிவப்பு, 4 கறுப்புக் குண்டுகள் உள்ள இரு பைகளிலிருந்து எடுக்கப்படும் இரண்டு குண்டுகளும் சிவப்பாக இருப்பதற்கான வாய்ப்பினைக் காண்போம்.

முதல் பையிலிருந்து எடுக்கப்படும் குண்டு சிவப்பாக இருப்பதற்கான நிகழ்திறன் $= \frac{6}{10}$

இரண்டாவது பையிலிருந்து எடுக்கப்படும் குண்டு சிவப்பாக இருப்பதற்கான நிகழ்திறன் $= \frac{6}{10}$

எடுக்கப்பட்ட இரு குண்டுகளும் சிவப்பாக இருப்பதற்கான நிகழ்திறன் $= \frac{6}{10} \times \frac{6}{10} = \frac{36}{10}$

$$p = \frac{9}{25}$$

நியதி 3: A_1, A_2 என்ற இரு சார்புடை நிகழ்ச்சிகள் சேர்ந்து நிகழ்வதற்கான நிகழ்திறன், A_1 என்ற நிகழ்ச்சி நிகழ்வதற்கான நிகழ்திறனையும், A_2 என்ற நிகழ்ச்சி நிகழ்வதற்கான நிபந்தனை நிகழ்திறனையும் (A_1 நிகழ்ந்தபின் A_2 நிகழ்வதற்கான நிகழ்திறன்) பெருக்கிவரும் மதிப்பிற்குச் சமமாகும். இதுவே A_2 என்ற நிகழ்ச்சி நிகழ்வதற்கான நிகழ்திறனையும், A_1 என்ற நிகழ்ச்சியின் நிபந்தனை நிகழ்திறனையும் (A_2 நிகழ்ந்தபின் A_1 நிகழ்வதற்கான நிகழ்திறன்) பெருக்கிவரும் மதிப்பிற்குச் சமமாகவும் கொள்ளலாம்.

$$P(A_1 A_2) = P(A_1) \cdot P(A_2/A_1)$$

$$P(A_1 A_2) = P(A_2)\, P(A_1/A_2)$$

A_1, A_2 இரண்டும் சார்பிலா (independent) நிகழ்ச்சிகளாயின்,

$P(A_1/A_2) = P(A_1)$, $P(A_2/A_1) = P(A_2)$ என அமைந்து,

$P(A_1 A_2) = P(A_1). P(A_2)$-க்குச் சமமாகிவிடும்.

காட்டு :

ஆறு சிவப்புப் பந்துகளும் நான்கு கறுப்புப் பந்துகளும் உள்ள பையிலிருந்து அடுத்தடுத்து எடுக்கப்படும் இரு பந்துகளும் (முதலில் எடுத்த பந்தைப் பையில் போடாமல் இரண்டாவது பந்தை எடுக்கும்போது) சிவப்பாக இருப்பதற்கான நிகழ்திறனைக் கணக்கிடுவோம்.

முதல் பந்து சிவப்பாக அமைவதற்கான நிகழ்திறன் $= \frac{6}{10}$

முதல் பந்து சிவப்பாக அமைந்து, எடுத்த பந்தைப் பையில் போடாமல் இருக்கும்போது இரண்டாவது பந்து சிவப்பாக அமைவதற்கான நிகழ்திறன் $= \frac{5}{9}$

அடுத்தடுத்து எடுக்கப்படும் இரு பந்துகள் சிவப்பாக இருப்பதற்கான நிகழ்திறன் $= \frac{6}{10} \times \frac{5}{9} = \frac{30}{90}$

காட்டு :

பத்துச் சிவப்பு, இருபது வெள்ளைப் பந்துகள் உள்ள பெட்டியிலிருந்து எடுக்கப்படும் முதல் பந்து சிவப்பாகவும், இரண்டாவது பந்து வெள்ளையாகவும் அமைவதற்கான (முதல் பந்தைப் பெட்டியில் போடாமல் இருக்கும்போது) நிகழ்திறனைக் காண்போம்.

சிவப்புப் பந்து எடுப்பதற்கான நிகழ்திறன் $= \frac{10}{30} = \frac{1}{3} = p(A_1)$

சிவப்புப் பந்து எடுத்தபின் அதைப் பெட்டியில் போடாமல் இருக்கும்போது வெள்ளைப் பந்து எடுப்பதற்கான நிகழ்திறன் $= \frac{20}{29} = p(A_2/A_1)$

அடுத்தடுத்து எடுக்கப்படும் இரு பந்துகளும் முறையே சிவப்பாகவும் வெள்ளையாகவும் அமைவதற்கான நிகழ்திறன் $= p(A_1) P(A_2/A_1) = \frac{1}{3} \times \frac{20}{29} = \frac{20}{78}$

காட்டு :

இரண்டு கறுப்பு, மூன்று வெள்ளை, நான்கு சிவப்புப் பந்துகள் உள்ள பையிலிருந்து அடுத்தடுத்து எடுக்கப்படும் இரண்டு பந்துகளில் முறையே முதல் பந்து சிவப்பாகவும் (முதல் பந்தைப் பையில் போடாமல் இருக்கும்போது) இரண்டாவது பந்து வெள்ளையாகவும் இருப்பதற்கான நிகழ்திறனைக் கணக்கிடுவோம்.

இக் கூட்டு நிகழ்ச்சி நிகழ்வதற்கான முறையை இரண்டு வழிகளில் முறைப்படுத்தலாம். முதல் பந்தின் நிறம்—இரண்டாவது பந்தின் நிறம். இதன் துணைகொண்டு அடியிற்காணும் பட்டியலை அமைக்கமுடியும்.

	B_1	B_2	B_3
A_1	2	6	8
A_2	6	6	12
A_3	8	12	12

A_1, A_2, A_3 என்ற நிகழ்ச்சிகள் முதல் பந்தானது முறையே கறுப்பு, வெள்ளை, மற்றும் சிவப்பாக இருப்பதைக் குறிக்கட்டும். B_1, B_2, B_3 என்பன அம் முறைப்படி அமைந்த இரண்டாவது பந்தின் நிறத்தைக் குறிக்கும்.

நிகழ்ச்சி நிகழ்வதற்கான மொத்த எண்ணிக்கை $= 9 \times 8 = 72$

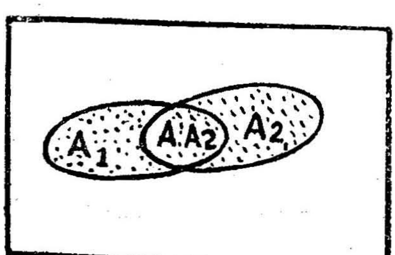
முதல் பந்து சிவப்பாகவும் இரண்டாவது பந்து வெள்ளையாகவும் நிகழ்வதற்கான சாதக நிகழ்ச்சிகள் $\Big\}$ $= 12.$ $A_3 B_2$

$$p = \frac{12}{72} = \frac{1}{6}$$

நியதி 4 : A_1, A_2 என்ற ஒன்றையொன்று விலக்கும் தன்மை பெறாத இரு நிகழ்ச்சிகளில் ஏதாவது ஒன்று நிகழ்வதற்கான நிகழ்திறன், ஒவ்வொரு நிகழ்ச்சியும் தனித்தனியே நிகழ்வதற்கான நிகழ்திறன்களின் கூட்டுத் தொகையிலிருந்து, இரு நிகழ்ச்சிகளும் சேர்ந்து நிகழ்வதற்கான நிகழ்திறனைக் கழித்துவரும் மதிப்பிற்குச் சமமாகும்.

$$p(A_1 + A_2) = p(A_1) + p(A_2) - p(A_1 A_2)$$

இந் நியதியை ஒரு தளத்தில் குறிக்கப்படும் நிகழ்திறப் புள்ளிகளால் எளிதில் உணரலாம்.



படம் 4.1

A_1 என்ற வட்டத்தில் உள்ள புள்ளிகள், A_1 என்ற நிகழ்ச்சிக்குச் சாதகமானவையாகவும், A_2 என்ற நிகழ்ச்சிக்குச் சாதகமான நிகழ்ச்சிகள் A_2 என்ற வட்டத்தில் உள்ள புள்ளிகளாகவும் கொண்டால், A_1 மற்றும் A_2 ஆகிய இரண்டில் ஏதாவதொன்று நிகழ்வதற்குச் சாதகமான புள்ளிகள் A_1, A_2 ஆகிய இரு வட்டங்களிலும் உள்ள மொத்தப் புள்ளிகளுக்கும் சமமாகும். இரு வட்டங்களிலும் உள்ள புள்ளிகளைக் கூட்டும்போது $(A_1 A_2)$ எனும் அடைப்பில் உள்ள புள்ளிகள் இருமுறை கூட்டப்படுவதைத் தெரிய வருகிறோம். எனவே, A_1, A_2 ஆகிய இரண்டில் ஏதாவதொன்று நிகழ்வதற்கான சாதக நிகழ்ச்சிகள்

$$(A_1) + (A_2) - (A_1 A_2) \text{ என அமைகிறது.}$$

எனவே, A_1, A_2 எனும் ஒன்றையொன்று விலக்கும் தன்மை பெறாத இரு நிகழ்ச்சிகளில் ஏதாவதொன்று நிகழ்வதற்கான நிகழ்திறன்

$$p(A_1 + A_2) = p(A_1) + p(A_2) - p(A_1 A_2)$$

காட்டு :

இரு பைகளில் பை ஒன்றுக்கு 6 சிவப்பு, 4 கறுப்புப் பந்துகள் உள்ளன. ஒவ்வொரு பையிலிருந்தும் ஒரு பந்து எடுக்கப்படும் போது எடுக்கப்பட்ட இரு பந்துகளில் ஒன்று சிவப்பாக இருப்பதற்கான நிகழ்திறனைக் கணக்கிடுவோம்.

முதல் பையிலிருந்து சிவப்புப் பந்து எடுப்பதற்கான நிகழ்திறன் $= p(A_1) = p_1 = \frac{6}{10}$

இரண்டாவது பையிலிருந்து சிவப்புப் பந்து எடுப்பதற்கான நிகழ்திறன் $= p(A_2) = p_2 = \frac{6}{10}$

இரண்டு பைகளிலிருந்தும் சேர்ந்தாற்போல் சிவப்புப் பந்து எடுப்பதற்கான நிகழ்திறன் $= \frac{6}{10} \cdot \frac{6}{10} \quad \frac{36}{100}$

ஒன்றையொன்று விலக்கும் தன்மையல்லாத இரு நிகழ்ச்சிகளில் ஏதாவதொன்று நிகழ்வதற்கான நிகழ்திறன் =

$$p(A_1 + A_2) = p(A_1) + p(A_2) - P(A_1 A_2)$$

$$= \frac{6}{10} + \frac{6}{10} - \frac{6}{10} \cdot \frac{6}{10} = \frac{21}{52}$$

காட்டு :

இரு பகடைகளை உருட்டுவதால் குறைந்தது 6 என்ற தொகை கிடைப்பதற்கான நிகழ்திறனைக் கணக்கிடுவோம்.

இரு பகடைகளை உருட்டுவதால் குறைந்தது 6 என்ற தொகை ஆறு, ஏழு, எட்டு, ஒன்பது, பத்து, பதினொன்று, பன்னிரண்டு ஆகிய தொகை கிடைப்பதற்கான நிகழ்திறன்களின் கூட்டுத் தொகைக்குச் சமமாகும்.

இருபகடைகளைச் சேர்ந்தாற்போல் உருட்டுவதால் ஏற்படும் மொத்த நிகழ்ச்சிகளின் எண்ணிக்கை $= 6 \times 6 = 36$.

இரு பகடைகளைச் சேர்ந்தாற்போல் உருட்டுவதால் குறைந்தது ஆறு என்ற தொகை கிடைப்பதற்கான சாதக நிகழ்ச்சிகள்.

ஆறு விழுவதற்கான சாதக நிகழ்ச்சிகள்:

முதற் பகடை	:	1	2	3	4	5	
இரண்டாம் பகடை	:	5	4	3	2	1	5

ஏழு விழுவதற்கான சாதக நிகழ்ச்சிகள்:

முதற் பகடை	:	1	2	3	4	5	6	
இரண்டாம் பகடை	:	6	5	4	3	2	1	6

எட்டு விழுவதற்கான சாதக நிகழ்ச்சிகள் :

முதற் பகடை	:	2	3	4	5	6	
இரண்டாவது பகடை	:	6	5	4	3	2	5

ஒன்பது விழுவதற்கான சாதக நிகழ்ச்சிகள் :

முதற் பகடை	:	3	4	5	6	
இரண்டாம் பகடை	:	6	5	4	3	4

பத்து விழுவதற்கான சாதக நிகழ்ச்சிகள் :

முதற் பகடை	:	4	5	6	
இரண்டாம் பகடை	:	6	5	4	3

பதினொன்று விழுவதற்கான சாதக நிகழ்ச்சிகள் :

முதற் பகடை	:	5	6	
இரண்டாம் பகடை	:	6	5	2

பன்னிரண்டு விழுவதற்கான சாதக நிகழ்ச்சிகள் :

முதற் பகடை	:	6	
இரண்டாம் பகடை	:	6	1

குறைந்தது ஆறு கிடைப்பதற்கான சாதக நிகழ்ச்சிகள் $= 5+6+5+4+3+2+1$

$= 26$

$$\left.\begin{matrix}\text{இரண்டு பகடைகளை உருட்டும்போது}\\ \text{குறைந்தது ஆறு கிடைப்பதற்}\\ \text{நிகழ்திறன்}\end{matrix}\right\} = \frac{\text{சாதக நிகழ்ச்சிகள்}}{\text{மொத்த நிகழ்ச்சிகள்}}$$

$$p = \frac{m}{n} = \frac{26}{36} = \frac{13}{18}$$

நிகழ்முன் நிகழ்திறன் மதிப்பிடும்போது, நிகழ்ச்சி நிகழக் கூடிய எல்லா வழிகளேயும் கணக்கிட வேண்டியவர்களாக உள்ளோம். இக் கணக்கீட்டு முறையைச் சேர்மானம் (combination), வரிசைமாற்றம் (permutation), ஈருறுப்புத் தேற்ற விரிவு (binomial expansion) ஆகியவற்றின் உதவியால் மிகவும் எளி தாக்கிக் கொள்ள முடியும். புள்ளியியல் முழுதும் நிகழ்திறனைச் சார்ந்துள்ளதால், ஒருவர் எந்த அளவிற்கு நிகழ்திறன்பற்றிச் சரியாகப் புரிந்துகொள்ள முடிகிறதோ அந்த அளவிற்குப் புள்ளியியல் நியதிகளேயும் சரிவர உணர்ந்துகொள்ள முடியும். நிகழ்திற நியதிகளைப் புரிந்துகொள்வதற்கும் கணக்குகளை எளிதாகச் செய்வதற்கும் அடியிற்காணும் நியதிகள் மிக முக்கியமானவையாக உள்ளன.

A என்ற நிகழ்ச்சி m-வழிகளிலும், B என்ற நிகழ்ச்சி n-வழிகளிலும் நிகழ்வதாகக் கொண்டால் A, A சேர்ந்து நிகழாதிருக்கும்போது A அல்லது A, '$m + n$' வழிகளில் நிகழும்.

'm' வழிகளில் நிகழும் A என்ற நிகழ்ச்சியும், 'n' வழிகளில் நிகழும் B என்ற நிகழ்ச்சியும் சேர்ந்து நிகழ்வதற்கான மொத்த வழிகள் '$m \times n$' ஆகும்.

A என்ற நிகழ்ச்சியைச் சீட்டுக் கட்டிலிருந்து எடுக்கப்படும் ஸ்பேட்டினுக்கும் (Spade), B என்ற நிகழ்ச்சியைச் சீட்டுக் கட்டிலிருந்து எடுக்கப்படும் டைமனுக்கும் (Diamond) ஒப்பிட்டால், இந் நிகழ்ச்சி ஒவ்வொன்றும் பதின்மூன்று வழிகளில் நடைபெறுவதைத் தெரியவருகிறோம். ஸ்பேடு அல்லது டைமன் கிடைக்கக் கூடிய மொத்த வழிகள் $13 + 13 = 26$. அடுத்தடுத்து எடுக்கப்படும் இரண்டு சீட்டுகள் டைமன் மற்றும் ஸ்பேடாக இருக்க, முதலில் எடுக்கப்படும் டைமன் (பதின்மூன்று டைமன் சீட்டுகளில் ஏதாவதொரு டைமன்) இரண்டாவது எடுக்கப்படும் ஸ்பேட்டுடன் (பதின்மூன்று ஸ்பேடுகளில் ஏதாவதொன்று) சேரக் கூடுமாதலால், இவ் விரு நிகழ்ச்சிகளும் சேர்ந்து நிகழ்வதற்கான மொத்த வழிகள் $13 \times 13 = 169$ என அமைகின்றன.

மேற்கண்ட இரு கோட்பாடுகளும் நிகழ்ச்சிகளின் எண்ணிக்கையை அதிகப்படுத்தினாலும், அதே முறையில் அமைந்து பொது விதியாகக் கொள்வதற்கு வழி செய்கிறது. A_1 A_2......A_n என்ற நிகழ்ச்சிகள் m_1, m_2,......m_n வழிகளில் நிகழ்வதாகக் கொண்டால், A_1 அல்லது A_2 அல்லது .....A_n ஆகியவற்றில் ஏதாவதொன்று நிகழ்வதற்கான மொத்த வழிகள் $m_1 + m_2 + \ldots + m_n$ எனவும், A_1 A_2 ... ... A_n ஆகியவை சேர்ந்து நிகழ்வதற்கான மொத்த வழிகள் $m_1 \times m_2$ ... m_n எனவும் அமைகின்றன.

வரிசைமாற்றம் (Permutation)

வேறுபாடுடைய 'n' பொருள்களை எடுத்துக்கொள்வோம். அவை 'n' வேறுபட்ட எண்களாகவோ அல்லது 'n' வேறுபட்ட நிறங்களைக்கொண்ட பொருள்களாகவோ அல்லது 'n' வேறுபட்ட எழுத்துகளாகவோ இருக்கட்டும். இக் கூட்டிலிருந்து 'r' பொருள்களை எடுத்து வரிசையாக அமைப்போம். இரு பொருள்கள் இருக்குமிடம் மாற்றப்படும்பொழுது வேறுவிதமான வரிசை மாற்றம் கிடைப்பதாகக் கொள்வோம். 'n' பொருள்களிலிருந்து 'r' பொருள்களை எடுத்தடுக்குவதால் கிடைக்கக்கூடிய மொத்த வரிசை மாற்றங்களை அடியிற்காணும் முறைப்படி கணக்கிடலாம். வரிசையின் முதலிடம் கொடுக்கப்பட்ட 'n' பொருள்களில் ஏதாவதொன்றாக இருக்கலாம். எனவே, வரிசையின் முதலிடம் 'n' முறைகளில் அமைகின்றன. முதலிடம் அமைக்கப்பட்டபின் வரிசையின் இரண்டாவது இடம் மீதமுள்ள $(n-1)$ பொருள்களில் ஏதாவதொன்றாக இருக்கலாம். எனவே, இரண்டாவது இடம் $(n-1)$ முறைகளில் அமைகின்றன. வரிசையின் முதல், இரண்டாவது இடங்கள் $n(n-1)$ முறைகளில் அமையும். இம் முறைப்படி தொடர்ந்தால் இறுதியாக உள்ள 'r' இடத்திற்கு மீதமுள்ள $(n-r+1)$ பொருள்களில் ஏதாவதொன்று அமையும். 'r'-வது இடம் $(n-r+1)$ முறைகளில் அமையும். எனவே 'r' இடங்களும் $n(n-1)(n-2)\ldots\ldots(n-r+1)$ முறைகளில் அமையும். 'n' பொருள்களிலிருந்து 'r' பொருள்களை எடுத்து அமைக்கும் முறையே வரிசைமாற்றம் எனப்படும். இதுவே கணக்கியல் குறியீடாக nPr (என் பி ஆர்) எனக் குறிக்கப்படுகிறது.

$$nPr = n(n-1)\ldots\ldots\ldots(n-r+1).$$

காட்டு :

a, b, c, d என்ற நான்கெழுத்துக் கூட்டிலிருந்து இரண்டெழுத்து எடுத்தடுக்குவதால் கிடைக்கக்கூடிய மொத்த வரிசைமாற்றங்கள் $4P_2 = 4(4-1) = 12$.

இவற்றை முறையே,

$$ab \quad ba \quad ac \quad ca \quad ad \quad da$$
$$bc \quad cb \quad bd \quad db \quad cd \quad dc$$

எனக் குறிப்பிடலாம். எழுத்துகளின் இருப்பிடத்திற்கு முக்கியத்துவம் தந்திருப்பது உணரத்தக்கது. ab, ba மற்றும் cd, dc ஆகியவை மாறுபட்ட வரிசைமாற்றங்களாகக் குறிக்கப்பட்டுள்ளமை இத் தன்மையைக் குறிக்கும்.

$r = n$ எனக் கொள்ளுமிடத்துக் கிடைக்கக்கூடிய வரிசை மாற்றங்கள் $n\,P_n$ எனக் குறிக்கப்படும்.

$n\,P_n = n\,(n-1)\ldots\ldots\ldots 3.2.1.$

இதுவே n பொருள்களிலிருந்து எடுக்கப்பட்ட n பொருள்களையும் n இடத்தில் அமைக்கும் வரிசைமாற்றங்களாகும். முதலிடம் 'n' முறைகளிலும், இரண்டாவது இடம் $(n-1)$ முறைகளிலும்......இறுதியாக உள்ள இடம் ஒரு வழியிலும் நிகழ்வதால், இதன் மொத்த வரிசைமாற்றங்கள் $n\,(n-1)\;(n-2)\ldots\ldots 3.2.1$ என அமைகின்றன. அடுத்தடுத்த முதல் 'n' இயற்கை எண்ணின் (natural numbers) பெருக்கற்பலனை இது குறிக்கும். இப் பெருக்குத்தொகை காரணீயப் பெருக்கம் (factorial) எனச் சொல்லப்படும். இது $\underline{|n}$, $n!$ என்ற குறியீடுகளால் குறிக்கப்பட்டு, 'ஃபேக்டோரியல் n' எனச் சுட்டப்படுகிறது.

$$n\,P_n = n\,(n-1)\,(n-2)\ldots\ldots 3.2.1 = n!$$

$$\therefore \quad 2! = 1 \times 2 = 2$$
$$3! = 1 \times 2 \times 3 = 6$$
$$\vdots$$
$$9! = 1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8 \times 9 = 362880$$

என்ற மதிப்புகளைத் தரும்.

'n' பூச்சியமாக அமையும்போது காரணீயப் பெருக்கத்தின் முறைகளை அமைப்பதற்காக $0! = 1$ எனக் கொள்கிறோம்.

சேர்மானம் (Combination)

எடுக்கப்பட்ட பொருள்களின் இருப்பிடத்திற்கு முக்கியத்துவம் தராமல் அமைக்கப்படும் வரிசைகள் சேர்மானம் எனப்படும். a, b, c, d என்ற கூட்டிலிருந்து எடுத்தமைக்கப்படும் ab என்ற சேர்மானம் ba என்ற சேர்மானத்திற்குச் சமமானதாகவும், ca என்ற சேர்மானத்திற்கு மாறுபட்டதாகவும் கொள்ளப்படும்.

'n' பொருள்களிலிருந்து 'r' பொருள்களை எடுத்தமைக்கப்படும் மொத்தச் சேர்மானங்கள் $\binom{n}{r}$ எனக் குறிக்கப்படும். இதுவே nC_r (என்சிஆர்) எனவும் குறிக்கப்படலாம். இக் குறியீடுகள் $\left[\binom{n}{r}, nC_r\right]$ 'n' பொருள்களிலிருந்து 'r' பொருள்களை எடுத்தமைப்பதால் கிடைக்கும் மொத்தச் சேர்மானங்களைக் குறிக்கும்.

மொத்த வரிசைமாற்றங்களையும், சேர்மானங்களையும் ஒப்பிட்டுச் சேர்மானங்களின் மதிப்பைக் கணக்கிடலாம். வரிசை மாற்றத்தில் 'r' பொருள்கள் எடுக்கப்பட்டுப் பொருள்களின் இருப்பிடத்திற்கு முக்கியத்துவம் கொடுத்து வரிசையாக அமைக்கப்படுகின்றன. ab என்ற வரிசைமாற்றமும் ba என்ற வரிசைமாற்றமும் மாறுபட்டதெனக் கொள்ளப்படும். முதல் வரிசைமாற்றத்தில் 'a' முதலிடத்திலும், இரண்டாவது வரிசைமாற்றத்தில் 'b' முதலிடத்திலும் அமைந்துள்ளன. சேர்மானத்தில் பொருள்கள் இருக்கும் இடங்களுக்கு முக்கியத்துவம் இல்லாததால் 'n' பொருள்களிலிருந்து 'r' பொருள்களை எடுக்கும் முறைகளை மட்டும் கணக்கிடுகிறோம். இந்த 'r' பொருள்கள் தம்முள் எப்படி அமைந்தாலும் ஒரே சேர்மானத்தைச் சேர்ந்ததாகக் கொள்ளப்படும். எனவே, வரிசைமாற்றம் என்பது சேர்மானத்தின் மதிப்பைச் சேர்மானத்தில் உள்ள பொருள்கள் தங்கட்குள் ஏற்படுத்திக் கொள்ளும் மாறுபட்ட அமைப்புகளின் முறையால் பெருக்கிவரும் தொகைக்குச் சமமாகிறது. a, b, c என்ற மூன்று பொருள்களிலிருந்து இரண்டு பொருள்கள் எடுப்பதற்கான வரிசைமாற்றத்தைக் காண்போம்.

$$3P2 = 3 \cdot 2 = 6$$
$$= ab, ba, bc, cb, ac, ca$$

a b, b a என்ற இரு வரிசைமாற்றங்களும் ஒரே சேர்மானமாகக் கொள்ளப்படும். இதேபோல் (bc, cb) (ac, ca) ஆகியவையும் ஒரே சேர்மானத்தைக் குறிப்பனவாகக் கொள்ளப்படும். எடுக்கப்பட்ட ஒவ்வொரு சோடி உறுப்பும் தம்முள் இரு வேறுபட்ட வரிசைமாற்றங்களையும், ஒரு சேர்மானத்தையும் கொடுக்கின்றன.

வரிசைமாற்றம் = (சேர்மானம்) × (சேர்மான உறுப்புகளின் வேறுபட்ட அமைப்பு முறைக்கான வழிகள்)

$$3P2 = \binom{3}{2}\ 2\,!.$$

$$\binom{3}{2} = \frac{3 \cdot 2}{1 \cdot 2} = \frac{6}{2} = 3$$

மூன்று பொருள்களிலிருந்து இரண்டு பொருள்களை எடுப்பதற்கான சேர்மானத்தின் மதிப்பு மூன்றாகும். இதுவே, பொதுமைப் படுத்தும்போது 'n' பொருள்களிலிருந்து r பொருள்களை எடுப்பதை $\binom{n}{r}$ எனவும், r பொருள்களை தங்கட்குள் வரிசைப் படுத்தும் முறைகளை $r!$ எனவும் கொள்கிறோம்.

$$nPr = nCr \cdot r!$$

$$nCr = \frac{nCr}{r!} = \frac{n\,(n-1)\,(n-2)\ldots\ldots(n-r+1)}{1\,.\,2\,.\,3\ldots\ldots\ldots\ldots r}$$

வலப் புறத்தில் அமைந்த மதிப்பை $(n-r)\ (n-r+1)\ \ldots\ldots 3\ .\ 2.\ 1$—ஆல் பெருக்கியும், வகுத்தும் அதன் மதிப்பை மாறுமல் செய்யும் போது

$$nCr = \frac{n(n-1)\ldots\ldots(n-r+1)}{1\,.\,2\,.\,3\ldots\ldots r} \times \frac{(n-r)\ (n-r-1)\ldots 3\,.\,2\,.\,1}{(n-r)\,(n-r-1)\ldots 3\,.\,2\cdot 1}$$

$$= \frac{n!}{r!\ (n-r)!}$$

காட்டு :

abcd என்ற நான்கெழுத்துச் சேர்க்கையிலிருந்து இரண்டு எழுத்துகள் கொண்ட சேர்மானங்களைக் கணக்கிடு.

$$\binom{4}{2} = \frac{4 \cdot 3}{1 \cdot 2} = \frac{12}{2} = 6.$$

இவை முறையே *ab, ac, ad, bc, bd, cd* என அமைகின்றன.

காட்டு :

40 நகரங்கள் உள்ள பட்டியலிலிருந்து 5 நகரங்கள் உள்ள ராண்டம் மாதிரிகள் (Random samples) எவ்வளவு கிடைக்கும் என்பதைக் கணக்கிடு.

40 நகரங்களிலிருந்து 5 நகரங்கள் கொண்ட கூறுகளின் எண்ணிக்கை, 40 பொருள்களிலிருந்து 5 பொருள்களை எடுக்கும் சேர்மானங்களின் எண்ணிக்கைக்குச் சமமாகும். ஒவ்வொரு கூறிலும் அமையும் நகரங்கள் மாறுபட்டதாகவும், அவற்றின் அமையும் இடத்திற்கு முக்கியத்துவம் இல்லாததாலும், இவை சேர்மானங்களின் எண்ணிக்கைக்குச் சமமானதாக அமைகின்றன.

$$\left.\begin{array}{r}\text{40 நகரங்களிலிருந்து 5 நகரங்கள்}\\ \text{கொண்ட கூறுகளின் எண்ணிக்கை}\end{array}\right\} = \binom{40}{5}$$

$$= \frac{40 \times 39 \times 38 \times 37 \times 36}{1\cdot 2\cdot 3\cdot 4\cdot 5}$$

$$= 658008.$$

காட்டு :

சீட்டுக் கட்டொன்றிலிருந்து எடுக்கப்பட்ட இரண்டு சீட்டு களில் ஒன்று ஸ்பேடாகவும் (spade), மற்றது கிளப்பாகவும் (club) இருப்பதற்கான நிகழ்திறன் யாது ?

நிகழ்ச்சிகள் நிகழும் மொத்த எண்ணிக்கை, 52 சீட்டு களிலிருந்து இரண்டு சீட்டுகள் எடுப்பதற்கான சேர்மானங்களைக் குறிக்கும். நிகழ்ச்சிகள் நிகழும் வரிசைக்கு முக்கியத்துவம் தராததால் இது சேர்மான வகையைக் குறிக்கும்.

$$\text{மொத்த வழிகள் : } \binom{52}{2} = \frac{52\cdot 51}{1\cdot 2} = 1326.$$

நிகழ்ச்சிக்குச் சாதகமானவை, எடுக்கப்படும் சீட்டு ஸ்பேடாகவோ கிளப்பாகவோ அமைவதற்கான வழிகளின் பெருக்குத் தொகைக்குச் சமமாகும்.

எடுக்கப்படும் சீட்டு ஸ்பேடாக இருப்பதற்கான சாதக நிகழ்ச்சி $= 13.$

எடுக்கப்படும் சீட்டு கிளப்பாக இருப்பதற்கான சாதக நிகழ்ச்சி $= 13.$

$$\left.\begin{array}{r}\text{எடுக்கப்பட்ட இரு சீட்டுகளும் ஸ்பேடு, கிளப்}\\ \text{பாக இருப்பதற்கான சாதக நிகழ்ச்சிகள்}\end{array}\right\} = 13 \times 13 = 169$$

$$\left.\begin{array}{r}\text{எடுக்கப்படும் இரண்டு சீட்டுகளில் ஒன்று}\\ \text{ஸ்பேடாகவும் மற்றது கிளப்பாகவும்}\\ \text{இருப்பதற்கான நிகழ்திறன்.}\end{array}\right\} = \frac{169}{1326} = \frac{13}{102}$$

காட்டு : 10, 15 நகரங்களாகப் பகுத்துள்ள 25 நகரங்களின் பட்டியலிலிருந்து 5 நகரங்களடங்கிய கூறினை எடுக்க வேண்டியவர் களாக உள்ளோம். பகுத்துள்ள முதல் கூறிலிருந்து இரண்டும், இரண்டாவது கூறிலிருந்து மூன்றுமாக எடுத்தால் கிடைக்கக் கூடிய மொத்தக் கூறுகளையும், பட்டியல் இரு பகுதிகளாகப் பிரிக்கப்படாமல் இருக்கும்போது கிடைக்கக் கூடிய மொத்தக் கூறுகளையும் கணக்கிடு.

பகுக்கப்படாமல் இருக்கும் 25 நகரங்களடங்கிய பட்டியலி லிருந்து 5 நகரங்கள் கொண்ட கூறுகள் எடுப்பது 25 பொருள் களிலிருந்து 5 பொருள்கள் கொண்ட சேர்மானங்களை அமைப் பதற்குச் சமமாகும்.

$$\binom{25}{5} = \frac{25\,!}{5\,!\cdot 20\,!} = 53130.$$

25 நகரங்கள் கொண்ட பட்டியலிலிருந்து 5 நகரங்கள் கொண்ட கூறு 53130 வழிகளில் எடுக்கலாம்.

10, 15 நகரங்களாகப் பிரிக்கப்பட்ட இரு பகுதிகளிலிருந்து முறையே 2, 3 நகரங்கள் எடுத்து 5 நகரங்கள் கொண்ட கூறுகள் எவ்வளவு அமையும் என்பதைக் கணக்கிடுவோம்.

10 நகரங்களிலிருந்து 2 நகரங்கள் எடுப்பது 10 பொருள் களிலிருந்து 2 பொருள்கள் எடுத்தமைக்கப்படும் சேர்மானங் களுக்குச் சமமாகும்.

$$\binom{10}{2} = \frac{10\cdot 9}{1\cdot 2} = \frac{90}{2} = 45.$$

15 நகரங்களிலிருந்து 3 நகரங்கள் எடுப்பது 15 பொருள் களிலிருந்து 3 பொருள்களை எடுத்தமைக்கப்படும் சேர்மானங் களுக்குச் சமமாகும்.

$$\binom{15}{3} = \frac{15\cdot 14\cdot 13}{1\cdot 2\cdot 3} = 455.$$

முதல் பகுதியிலிருந்து எடுக்கப்பட்ட 45 கூறுகளும் இரண் டாவது பகுதியிலிருந்து எடுக்கப்பட்ட 455 கூறுகளுள் ஏதாவ தொன்றுடன் சேர்த்து ஒவ்வொரு கூறின் உறுப்புகளும் 5 என அமையச் செய்யும். எனவே, 10 நகரங்களிலிருந்து 2 நகரங்கள் எடுத்த முறைகளை 15 நகரங்களிலிருந்து 3 நகரங்கள் எடுத்த முறைகளால் பெருக்குவதால் 10, 15 என இரு பகுதிகளாகப் பிரிக்கப்பட்ட 25 நகரங்களில் முறையே 2, 3 நகரங்கள் எடுத் தமைத்த 5 நகரங்கள் கொண்ட கூறுகளின் எண்ணிக்கை கிடைக்கும்.

10, 15 என்ற இரு பகுதிகளில் 2, 3 நகரங்கள் எடுக்கும்போது 5 நகரங்கள் கொண்ட கூறுகளின் எண்ணிக்கை $\Big\} = 45 \times 455$

$= 20475.$

காட்டு :

சீட்டுக் கட்டொன்றிலிருந்து எடுக்கப்பட்ட 5 சீட்டுகளில் எண் '3' பொறிக்கப்பட்டுள்ள இரு சீட்டுகள் இருப்பதற்கான நிகழ்திறனைக் காண்போம்.

சீட்டுக் கட்டிலுள்ள 4 மூன்றுகளையும் A_1 என்றும், மற்றச் சீட்டுகளை A_2 எனவும் கொள்வோம். நிபந்தனை நிகழ்திறனைக் கணக்கிடும்போது ஐந்து சீட்டுகளையும் ஒவ்வொன்றாக எடுப்பதாகவும் அவை ஒரு குறிப்பிட்ட முறையில் உதாரணமாக A_1 A_1 A_2 A_2 A_2 என நிகழ்வதாகவும் கொள்வோம்.

$$P(A_1 A_2 \ldots\ldots A_n) = P(A_1)\, P(A_2/A_1) P(A_3/A_1A_2) \ldots\ldots P(A_n/A_1 A_2 \ldots\ldots A_{n-1}).$$

$$P(A_1\, A_1\, A_2\, A_2\, A_2) = P(A_1)\, P(A_1/A_1)\, P(A_2/A_1\, A_1)\; P(A_2/A_1 A_1 A_2) \times P(A_2/A_1\, A_1\, A_2\, A_2).$$

$P(A_1) = \frac{4}{52}$ ஒரு மூன்று எடுப்பதற்கான நிகழ்திறன்.

$P(A_1/A_1) = \frac{3}{51}$ ஒரு மூன்றெடுத்த பின் இரண்டாவது மூன்று எடுப்பதற்கான நிகழ்திறன்.

$P(A_2/A_1A_1) = \frac{48}{50}$ இரண்டு மூன்றுகள் எடுத்தபின் மற்ற ஒரு சீட்டு எடுப்பதற்கான நிகழ்திறன்.

$P(A_2/A_1\, A_1\, A_2) = \frac{47}{49}$

$P(A_2/A_1\, A_1\, A_2\, A_2) = \frac{46}{48}$

$$P(A_1\, A_1\, A_2\, A_2\, A_2) = \frac{4}{52} \cdot \frac{3}{51} \cdot \frac{48}{50} \cdot \frac{47}{49} \cdot \frac{46}{48} = 0{\cdot}00399$$

இது ஒரு குறிப்பிட்ட முறையில் நடைபெற்ற நிகழ்ச்சிக்கான நிகழ்திறனாகும். நிகழ்ச்சி நிகழக்கூடிய எல்லா வழிகளையும் கணக்கிடவேண்டி உள்ளது. எடுக்கப்பட்ட 5 சீட்டுகளில் 2 சீட்டுகள் ஓர் இனத்தைச் சேர்ந்தனவாகவும் மற்ற மூன்று சீட்டு மற்றோர் இனத்தைச் சேர்ந்தனவாகவும் கொள்கிறோம்.

இவை செயல்படும் மொத்த வழிகள் : $\frac{5!}{3!\cdot 2!} = 10.$

எடுக்கப்பட்ட 5 சீட்டுகளில் இரண்டு மூன்றுகள் இருப்பதற்கான நிகழ்திறன் $= 10 \times 0{\cdot}00399$

$= 0{\cdot}0399.$

காட்டு :

8 கோப்பைத் தேநீரிலிருந்து, ஒரு குறிப்பிட்ட முறைப்படி தயார்செய்யப்பட்ட 4 தேநீர்க் கோப்பைகளைக் குறிப்பிடுவது

நமது ஆய்வாக இருக்கட்டும். இதற்கான நிகழ்திறனைக் கணக்கிடுவோம்.

இந் நிகழ்ச்சி நிகழ்வதற்கான மொத்த வழிகள் 8 கோப்பைகளிலிருந்து 4 கோப்பைகள் எடுப்பதற்கான சேர்மானங்களின் எண்ணிக்கைக்குச் சமமாகும்.

$$\binom{8}{4} = \frac{8\,!}{4!.4!} = 70$$

இந் நிகழ்ச்சிக்குச் சாதகமான நிகழ்ச்சி ஒன்றுதான் உள்ளது. 70 சேர்மானங்களில் ஓர் அமைப்பு மட்டும் குறிப்பிட்ட முறைப்படி தயார் செய்யப்பட்ட 4 தேநீர்க் கோப்பைகளைப் பெற்றிருக்கும்.

$$\text{நிகழ்திறன்} : \frac{\text{சாதக நிகழ்ச்சிகள்}}{\text{மொத்த நிகழ்ச்சிகள்}} = \frac{1}{70}$$

ஈருறுப்புத் தேற்ற விரிவு (Binomial Expansion)

$\binom{n}{r}$ என்ற சேர்மானத்தில், $n = 1, 2, 3, 4$, $r = 0, 1, 2, 3, 4$ என்ற மதிப்புகளைப் பெறும்போது கிடைக்கக்கூடிய சேர்மானங்களை அடியிற்காணும் அட்டவணையில் அமைக்கலாம்.

r \ n	1	2	3	4
0	1	1	1	1
1	1	2	3	4
2		1	3	6
3			1	4
4				1
	2	4	8	16

இவ் வட்டவணையில் கொடுக்கப்பட்டுள்ள சேர்மான எண்ணிக்கைகள் n அடுக்கு ஈருறுப்புத் தேற்ற விரிவு உறுப்புகளின் குணகங்களாக அமைகின்றன. $(a+b)^n$ என்ற 'n' அடுக்கு ஈருறுப்புத் தேற்ற விரிவு, $n = 1, 2, 3, 4$ என அமையும்போது அடியிற்காணும் உறுப்புகளைப் பெறும்.

குணகங்களின் கூட்டுத்தொகை

$(a+b) = 1 \cdot a + 1 \cdot b$	2
$(a+b)^2 = 1a^2 + 2ab + 1 \cdot b^2$	4
$(a+b)^3 = 1a^3 + 3a^2 b + 3ab^2 + 1b^3$	8
$(a+b)^4 = 1a^4 + 3a^3 . b + 3a^2 b^2 + 4ab^3 + 1 \cdot b^4$	16

ஈருறுப்புத் தேற்ற விரிவின் உறுப்புகளின் குணகமும், சேர்மான எண்ணிக்கைகளும் சமமாக இருப்பதைக் காணலாம். n-ன் மற்ற மதிப்புகளுக்கும் இது பொருந்தும்.

ஒரு சோடி நாணயங்கள் 'n' முறையில் சுண்டப்படுகின்றன. இச் சோதனையில் 'x' முறைகளில் இரு தலைகள் கிடைக்கக்கூடிய நிகழ்திறனைக் கணக்கிடுவோம்.

நமக்குத் தேவையான நிகழ்திறன் ஒரு குறிப்பிட்ட முறையில் நிகழ்வதாகக் கொள்வோம்.

$\frac{x}{S.S. \ldots S}\ \frac{n-x}{f f f. \ldots F}$, S என்பது இரு தலைகள் சேர்ந்து விழுவதையும் (வெற்றியையும்), F, இரு தலைகள் சேர்ந்து விழாமையையும் குறிக்கும். F–பூ விழுந்து தோல்வியைக் குறிக்கும். சார்பற்ற தன்மைத்தான நிகழ்ச்சிகளுள் இந் நிகழ்ச்சி நிகழ்வதற்கான நிகழ்திறன் $p^x\ q^{n-x}$. p என்பது இரு தலைகள் சேர்ந்து விழுவதற்கான நிகழ்திறனையும், q இரு தலைகள் சேர்ந்து விழாமைக்கான நிகழ்திறனையும் குறிக்கட்டும். இந் நிகழ்ச்சி நிகழ்வதற்கான மொத்த முறைகள் 'n' நிகழ்ச்சிகளிலிருந்து x நிகழ்ச்சிகளை எடுத்தமைக்கும் சேர்மானங்களுக்குச் சமமாகும். $p^x\ q^{n-x}$ ஐ $\binom{n}{x}$ ஆல் பெருக்கும்போது n நிகழ்ச்சிகளில் x வெற்றிகள் கிடைப்பதற்கான நிகழ்திறன் கிடைக்கும்.

$$(p\ x) = \binom{n}{x} p^x\ q^{n-x}.$$

காட்டு : ஒரு சோடி நாணயங்களை 200 முறை சுண்டுவதால் 60 முறை இரு தலைகள் கிடைப்பதற்கான நிகழ்திறனைக் கணக்கிடு.

ஒரு சோடி நாணயத்தைச் சுண்டுவதால் கிடைக்கக்கூடிய மொத்த நிகழ்ச்சிகளின் தன்மை

H H
H T
T H
T T

என 4 வழிகளில் அமையும். இதில் H H என அமையும் ஒரு முறை தான் நமக்குச் சாதகமானதாகும்.

ஒரு சோடி நாணயத்தைச் சுண்டுவதால் இரு தலைகள் கிடைப்பதற்கான நிகழ்திறன் $\frac{1}{4} = p$.

$$q = 1 - p = 1 - \tfrac{1}{4} = \tfrac{3}{4}.$$

ஒரு சோடி நாணயங்கள் 200 முறை சுண்டப்படுகின்றன. இதில் 60 முறை இரு தலைகள் விழுவதற்கான வழிகள், 200பொருள்களிலிருந்து 60 பொருள்களை எடுக்கும் சேர்மானத்திற்குச் சமமாகும் $\binom{200}{60}$.

$$\text{நிகழ்திறன்} = \binom{200}{60}\left(\frac{1}{4}\right)^{60}\left(\frac{3}{4}\right)^{200-60}$$

$$= \frac{200!}{60!\ 140!}(\cdot 25)^{60}(\cdot 75)^{140}$$

பயிற்சி :

குழந்தைகளின் மூன்று தொகுப்புகள் முறையே மூன்று பெண்கள், ஒரு பையன்; இரண்டு பெண்கள், இரண்டு பையன்கள்; ஒரு பெண், மூன்று பையன்கள் என உள்ளன. ஒவ்வொரு தொகுப்பிலிருந்தும் ஒருவரை எடுத்தமைக்கப்பட்ட கூட்டமொன்றில் உள்ள மூவரில் ஒரு பெண் இரு பையன்கள் இருப்பதற்கான நிகழ்திறனைக் கணக்கிடு.

2. ராண்டமாக எடுக்கப்பட்ட ஒரு லீப் வருடத்தில் 53 ஞாயிற்றுக் கிழமைகள் அமைவதற்கான நிகழ்திறனைக் கணக்கிடு.

3. ஆண் குழந்தை பிறப்பதற்கான நிகழ்திறனை '$\frac{1}{2}$' எனக் கொண்டு ஆறு குழந்தைகள் உள்ள குடும்பத்தில் அடியிற்காணும் நிகழ்ச்சிகள் நிகழ்வதற்கான நிகழ்திறனைக் கணக்கிடு.

(a) எல்லோரும் ஒரே பாலினைச் சேர்ந்திருப்பது.

(b) நான்கு மூத்த குழந்தைகள் ஆண், இறுதி இரண்டு குழந்தைகளும் பெண்.

(c) 3 குழந்தைகள் ஆண்.

4. 40 நல்ல நிலையிலும், 10 பழுதானவையுமான விதைகள் உள்ள கூட்டமொன்றிலிருந்து எடுக்கப்பட்ட 10 விதைகளில், எல்லாம் பழுதில்லாமல் இருப்பதற்கான நிகழ்திறனைக் கணக்கிடு.

5. இரண்டு வெள்ளை, மூன்று கறுப்புப் பந்துகள் உள்ள தட்டிலிருந்து இரண்டு பந்துகள் அடுத்தடுத்து எடுக்கப்படுகின்றன. முதல் பந்து வெள்ளையாகவும், இரண்டாவது பந்து கறுப்பாகவும் இருப்பதற்கான நிகழ்திறன் யாது? முதல் பந்தைத் தட்டில் போடாமல் இருக்கும்போது இந் நிகழ்ச்சி நிகழ்வதற்கான நிகழ்திறன் யாது?

6. சீட்டுக் கட்டொன்றிலிருந்து அடுத்தடுத்து எடுக்கப்படும் இரு சீட்டுகள் அடியிற் கண்டவாறு அமைவதற்கான நிகழ்திறனைக் காண்.

(a) இரண்டும் ஸ்பேடாக இருப்பதற்கான நிகழ்திறன்.

(b) இரண்டில் ஏதாவதொன்று ஸ்பேடாக இருப்பதற்கான நிகழ்திறன்.

(c) முதல் சீட்டினைக் கட்டில் போடாதிருக்கும்போது இரண்டும் ஸ்பேடாக இருப்பதற்கான நிகழ்திறன்.

7. மெண்டலியன் கூற்றுப்படி பட்டாணிகள் சமமான (smooth) உருவுடையனவாகவோ, நெளிவு வளைவு (wrinkled) உடையனவாகவோ இருக்கக்கூடும். சமமான பட்டாணிகள், நெளிவு வளைவுப் பட்டாணிகளைக் காட்டிலும் அதிகமானவை (dominant). (S S), (S W) சமனுடைப் பட்டாணிகளையும், (W W) நெளிவு வளைவுப் பட்டாணிகளையும் குறிக்கும்.

(Y, G), (S, W) பட்டாணிகளை (G, S), (W, W) பட்டாணிகளுடன் சேர்த்தால், கிடைக்கக்கூடிய பட்டாணிகளின் தன்மையையும் அதற்கான நிகழ்திறனையும் கணக்கிடு. (Y,G), (S,W)ஐ (G,G), (S'W) உடன் சேர்ப்பதாலும், (Y,G), (S,W)ஐ (Y,G). (S,W) வுடன் சேர்ப்பதாலும் கிடைக்கக்கூடிய வகைகளையும் நிகழ்திறனையும் கணக்கிடு.

8. மூன்று பகடைகளை வீசுவதால் கிடைக்கக்கூடிய 5 என்ற கூட்டுத்தொகைக்கான நிகழ்திறனைக் காண். ஆறு பகடைகளை வீசுவதால் எல்லா எண்களும் கிடைப்பதற்கான நிகழ்திறனைக் கணக்கிடு.

9. விதைப் பொட்டலங்கள் கொண்ட ஒரு முழுமைத் தொகுதியில் 10-ல் 1 பாகம் பழுதடைந்ததாகக் கொண்டு, 20 விதைகள் கொண்ட பொட்டலம் எடுத்ததில் (a) 2 விதைகள் பழுதடைந்ததாக வருவதற்கான நிகழ்திறனையும் (b) மூன்றும் அதற்கு மேலும் பழுதான விதைகள் இருப்பதற்குமான நிகழ் திறனையும் கணக்கிடு.

10. கீழே கொடுக்கப்பட்டுள்ள 'வாழ்க்கைப் பட்டியல்' (Life table) 1941-50, மக்கள்கணக்கெடுப்பிலிருந்து குறிக்கப் பட்டதாகும். *a*, *b*, *c* என்ற மூவர் 80, 84, 87 வயதுடையவர் களாகக் குறிக்கப்பட்டுள்ளனர். (1) மூவரும் இரண்டாண்டு களுக்கு உயிருடன் இருக்கும் நிகழ்திறன், (2) மூவரும் இரண் டாண்டில் இறப்பதற்கான நிகழ்திறன், (3) ஒருவராவது இரண் டாண்டிற்கு உயிருடன் இருப்பதற்கான நிகழ்திறன் ஆகிய வற்றின் மதிப்பைக் கணக்கிடு.

வயது x	உயிருடன் இருப்போர் lx	வயது x	உயிருடன் இருப்போர் lx
80	3213	86	749
81	2614	87	554
82	2098	88	402
83	1660	89	285
84	1294	90	198
85	993		

5. ஈருறுப்பு மற்றும் பாய்ஸான் பரவல்கள்

(Binomial and Poisson Distributions)

ஈருறுப்புப் பரவல் (Binomial distribution)

ஈருறுப்புத் தேற்றத்தின் அடிப்படைக் கருத்துகள் நிகழ்திறன் பகுதியில் விரிவாக்கப்பட்டதை உணர்வோம். தனித்த மாறிகளின் (discrete variables) பொதுப் பண்பைப் படிப்பதற்கு ஈருறுப்புப் பரவல் இன்றியமையாததாகும். முழுமைத் தொகுதியில் உள்ள உறுப்புகளில் நாம் குறிப்பிடும் பண்பு பெற்றவற்றை வெற்றியெனவும், பெருதவற்றைத் தோல்வியெனவும் கொள்கிறோம். பிறக்கும் குழந்தை ஆணுக இருப்பின் வெற்றிக்கும், பெண்ணுக இருப்பின் தோல்விக்கும் அல்லது கோதுமைக் கட்டொன்றில் உள்ள கதிர்களில், கூலவகைகளில் கதிரலகுள்ளவற்றை (awned) வெற்றிக்கும், கூலவகைகளில் கதிரலகற்றவற்றைத் (awnless) தோல்விக்கும் அல்லது ஒரு நாணயத்தைச் சுண்டும்போது தலே விழுவதை வெற்றிக்கும், பூ விழுவதைத் தோல்விக்கும் ஒப்பிடலாம்.

பிறக்க உள்ள இரு குழந்தைகளின் பாலினத் தெரியக் குறிக்கும் முறைகளில், அடியிற்காணும் ஈருறுப்புப் பரவல் கிடைப்பதைக் காணலாம்.

சோதனையின் விளைவு	நிகழ்திறன்	வெற்றிகள்	நிகழ்திறன்
S S	$p\ p$	2	p^2
S F	$p\ q$	1	
F S	$q\ p$	1	$2pq$
F F	$q\ q$	0	q^2

பிறந்த இரு குழந்தைகளுமே ஆணுகவும் (S, S), முதல் குழந்தை ஆணுகவும் இரண்டாவது குழந்தை பெண்ணுகவும் (S,F), முதல் குழந்தை பெண்ணுகவும் இரண்டாவது ஆணுகவும் (F, S), இரண்டுமே பெண்ணுகவும் (F, F) இருப்பது விளைவுகளின் மொத்த வழிகளாகும். பிறந்த முதற்குழந்தையின் பால் (ஆண், பெண்) இரண்டாவது குழந்தையின் பாலினைப் (sex) பாதிக்காத படி அமைந்துள்ளதால், ஈருறுப்புப் பரவலின் நியதிப்படி நிகழ் திறன் கணக்கிட ஏதுவாய் உள்ளது. வெற்றிகளைக் கணக்கிடும் போது S, F (ஆண், பெண்), F, S (பெண், ஆண்) ஆகிய இரு நிகழ்ச்சிகள் ஒவ்வொன்றிலும் ஒரு வெற்றி உள்ளதால், ஒரு வெற்றி உண்டாவதற்கான நிகழ்திறன் $2pq$ என அமைகிறது. இந் நிகழ்ச்சிகளின் மொத்த வாய்ப்பு ஒன்ருக முடிவதைக் காணலாம்.

$$p^2 + 2pq + q^2 = (p + q)^2 = 1^2 = 1$$

மூன்று குழந்தைகளைக் கொண்டு பாலின் (sex) அடிப்படையில் பிரிக்கும்போது கிடைக்கும் நிகழ்வெண் பரவல் அடியிற் கண்டவாறு அமையும்.

சோதனையின் விளைவு	நிகழ்திறன்	வெற்றிகள்	நிகழ்திறன்
S S S	$p\ p\ p$	3	p^3
S S F	$p\ p\ q$		
S F S	$p\ q\ p$	2	$3p^2 q$
F S S	$q\ p\ p$		
S F F	$p\ q\ q$		
F S F	$q\ p\ q$	1	$3\ pq^2$
F F S	$q\ q\ p$		
F F S	$q\ q\ p$	0	q^3

$n = 3$ என அமையும்போது ஏற்படக்கூடிய எட்டு வழிகளையும் அதற்கான நிகழ்திறன்களையும் பட்டியல் தருகிறது. 'n'-எம் மதிப்பைப் பெறினும் நிகழ்ச்சிகளுக்கான மொத்த வழிகளையும் அவற்றின் நிகழ்திறன்களையும் கொடுப்பதே ஈருறுப்புப் பரவலின் பொது வாய்பாடாகும்.

'x' வெற்றிகள் கொடுக்கக்கூடிய 'n' நிகழ்ச்சிகளை (x) வெற்றிகள்கொண்ட கூட்டாகவும், ($n-x$) தோல்விகள்கொண்ட கூட்டாகவும் இரு பிரிவுகளாகக் கொள்ளலாம். x வெற்றிகள், ($n-x$) தோல்விகள் நிகழ்வதற்கான நிகழ்திறன் $p^x q^{n-x}$ ஆனால், 'n' நிகழ்ச்சிகளில் 'x' வெற்றிகள் $\binom{n}{x}$ வழிகளில் நிகழும். நிகழ்திறன்மதிப்பை அந் நிகழ்ச்சி நிகழ்வதற்கான மொத்த வழிகளால் பெருக்கினால் கிடைக்கும் மதிப்பே ஒரு நிகழ்ச்சி நிகழ்வதற்கான நிகழ்திறனாகும். 'n' நிகழ்ச்சிகளில் x வெற்றிகளுக்கான நிகழ்திறன் மதிப்பு $p^x q^{n-x}$ ஆகவும், அதன் குணகம் $\binom{n}{x}$ ஆகவும் அமைவதால், n நிகழ்ச்சிகளில் x வெற்றிகள் கிடைப்பதற்கான நிகழ்திறன்

$$p(x) = \binom{n}{x} p^x q^{n-x}$$

$$= \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

இந் நிகழ்திறன் $(p+q)^n$ என்ற ஈருறுப்புக் கோவையின் (Binomial expansion) ஓர் உறுப்பாகும். இவ் ஈருறுப்புக் கோவையை விரிவுபடுத்துவதால் கிடைக்கும் அடுத்தடுத்த உறுப்புகளே $0, 1, 2 \ldots\ldots x, x+1, \ldots\ldots n$ வெற்றிகளுக்கான நிகழ்திறனாகும்.

காட்டு :

ஒரு பகடையை 5 தடவை உருட்டும்போது இரண்டு விழுவதற்கான நிகழ்திறனைக் கணக்கிடுவோம்.

வெற்றிக்கான வாய்ப்பு $= p = \frac{1}{6}$

$$q = 1-p = 1 - \frac{1}{6} = \frac{5}{6}$$

$$n = 5$$

$$p(x) = \binom{n}{x} p^x q^{n-x}$$

$$= \binom{5}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{5-2}$$

$$= 0.16$$

∴ இரண்டு கிடைப்பதற்கான நிகழ்திறன் $= 0{\cdot}16$

நிகழ்ச்சிகளின் எண்ணிக்கை குறைவாக அமையும்போது, பாஸ்கலின் முக்கோண அமைப்பிலிருந்து ஈருறுப்புக் குணகங்களை அடையலாம்.

பாஸ்கல் முக்கோணத்தில் (Pascal Triangle) அமையும் ஈருறுப்புக் குணகங்கள்

நிகழ்ச்சிகளின் எண்ணிக்கை n	ஈருறுப்புக் குணகங்கள்
	1
1	1 1
2	1 2 1
3	1 3 3 1
4	1 4 6 4 1
5	1 5 10 10 5 1
6	1 6 15 20 15 6 1
7	1 7 21 35 35 21 7 1
8	1 8 28 56 70 56 28 8 1

ஒவ்வொரு குணகமும் அதற்கு மேலுள்ள இட, வலப் புறத்தில் அமைந்த குணகங்களின் கூட்டுத்தொகைக்குச் சமமாகும். $n = 6$ என உள்ளபோது 20 என்ற குணகம் அதற்கு மேலுள்ள இரண்டு குணகங்களின் ($n = 5$, $10 + 10$) கூட்டுத்

தொகைக்குச் சமமாகும். 'n'-ன் எல்லா மதிப்பிற்குமான குணகங்கள் மத்தியிலிருந்து சமச்சீராக அமைந்துள்ளதைக் காணலாம்.

'n' நிகழ்ச்சிகளை ஒரு சோதனையாகக்கொண்டு 'N' முறைகள் இச் சோதனையைத் திரும்பச் செய்தால் $N.\binom{n}{x} p^{x} q^{n-x}$ தடவைகள் 'x' வெற்றிகள் கிடைக்குமென எதிர்பார்க்கிறோம். இதுவே ஈருறுப்பு நிகழ்வெண் பரவலாகும். ஈருறுப்பு நிகழ்வெண் பரவலைக் கீழ்க்கண்டவாறு அமைத்து அதன் கூட்டுச் சராசரியையும் திட்டவிலக்கத்தையும் கணக்கிடலாம்.

வெற்றிகளின் எண்ணிக்கை	வெற்றிக்குரிய நிகழ்திறன்	நிகழ்வெண்
0	$\binom{0}{n} q^{n}$	$N : q^{n}$
1	$\binom{n}{1} q^{n-1} p$	$N\binom{n}{1} q^{n-1} p$
2	$\binom{n}{2} q^{n-2} p^{2}$	$N\binom{n}{2} q^{n-2} p^{2}$
.	.	.
.	.	.
.	.	.
.	.	.
n	$\binom{n}{n} p^{n}$	$N\binom{n}{n} p^{n}$
மொத்தம்	$(q+p)^{n} = 1$	$N(q+p)^{n} = N$

இந் நிகழ்வெண் பரவலின் கூட்டுச்சராசரி 'np' ஆகவும், திட்டவிலக்கம் $\sqrt{npq}$ ஆகவும் அமைவதைக் கணக்கியல் முறைப்படி கணக்கிடலாம்.

கூட்டுச் சராசரி : $\bar{x} = np$

திட்ட விலக்கம் : $\sigma = \sqrt{npq}$

காட்டு :

10 விதைகள்கொண்ட 80 கூறுகளை விதைத்து, ஒவ்வொரு கூறிலும் முளைத்த செடிகளின் எண்ணிக்கையைக்கொண்டு கீழ்க் கண்ட பரவல் அமைக்கப்பட்டுள்ளது.

முளைத்த செடிகள் X வெற்றிகள்	0	1	2	3	4	5	6	7	8	9	10	மொத்தம்
முளைத்த கூறுகள் வெற்றிகளின் எண்ணிக்கை	6	20	28	12	8	6	0	0	0	0	0	80

10 விதைகள்கொண்ட 80 கூறுகளில், 6 கூறுகளில் ஒரு விதையுமே முளைக்காமலும், 20 கூறுகளில் ஒரு விதையும், 28 கூறுகளில் இரண்டு விதைகளும்.........6, 7, 8, 9, 10 விதைகள் முளைத்த கூறுகள் ஒன்றுமே அமையாததையும் சோதனை முடிவுகள் காட்டுகின்றன.

விதைகள் முளைப்பதை வெற்றியாகக்கொண்டு, ஒரு விதை முளைப்பது இன்னொரு விதை முளைப்பதைப் பாதிக்காமலிருப்பதால் ஈருறுப்புப் பரவலின் அடிப்படையில் கருதுகோள் நிகழ்வெண்களைக் (theoritical frequencies) கணக்கிடுவோம். நிகழ்வெண் பரவலுக்கான கூட்டுச்சராசரியைக் ($\bar{x}$) கணக்கிட்டு நிகழ்திறனைக் $\left(p = \frac{x}{n}\right)$ கணக்கிடுவோம்.

x_i	f_i	$x_i f_i$
0	6	0
1	20	20
2	28	56
3	12	36
4	8	32
5	6	30
6	0	.
7	0	.
8	0	.
9	0	.
10	0	.
மொத்தம்	80	174

$$\bar{x} = np = \frac{174}{80}$$

$$p = \frac{\bar{x}}{n} = 0{\cdot}2175.$$

$$q = 1 - p$$
$$= 0{\cdot}7825$$

ஈருறுப்பு நிகழ்வெண் பரவல் = $80(0{\cdot}7825 + 0{\cdot}2175)^{10}$

ஈருறுப்பு நிகழ்வெண் பரவலை விரித்துவரும் உறுப்புகளின் மதிப்பே 0, 1, 2...10 விதைகள் முளைப்பதற்கான கூறுகளின் எண்ணிக்கையாகும். இவ் வெண்களே கருதுகோள் நிகழ்வெண் களாகும். அடுத்த பக்கத்தில் கொடுக்கப்பட்டுள்ள அட்டவணை கருதுகோள் நிகழ்வெண்களைக் காணும் முறையை விளக்கமாகக் காட்டுகிறது.

முளைத்த விதைகளின் எண்ணிக்கை	$80 \times \binom{10}{x} (\cdot 2175)^x (\cdot 7825)^{10-x}$	கருதுகோள் நிகழ்வெண்	சோதனை முடிவு
0	$80 \times (0{\cdot}7825)^{10}$	6·89	6
1	$80 \times (0{\cdot}7825)^{9} \times (0{\cdot}2175) \times 10$	19·14	20
2	$80 \times (0{\cdot}7825)^{8} \times (0{\cdot}2175)^{2} \times 45$	23·95	28
3	$80 \times (0{\cdot}7825)^{7} \times (0{\cdot}2175)^{3} \times 120$	17·75	12
4	$80 \times (0{\cdot}7825)^{6} \times (0{\cdot}2175)^{4} \times 210$	8·64	8
5	$80 \times (0{\cdot}7825)^{5} \times (0{\cdot}2175)^{5} \times 252$	2·88	6
6	$80 \times (0{\cdot}7825)^{4} \times (0{\cdot}2175)^{6} \times 210$	0·67	0
7	$80 \times (0{\cdot}7825)^{3} \times (0{\cdot}2175)^{7} \times 120$	0·11	0
8	$80 \times (0{\cdot}7825)^{2} \times (0{\cdot}2175)^{8} \times 45$	0·01	0
9	$80 \times (0{\cdot}7825) \times (0{\cdot}2175)^{9} \times 10$	0·00	0
10	$80 \times (0{\cdot}2175)^{10}$	0·00	0
	மொத்தம்	80·04	80

கருதுகோள் நிகழ்வெண்களுக்கும், சோதனை முடிவுகளுக்கும் உள்ள வித்தியாசம் குறிப்பிடும்படி உள்ளதா (significant) அல்லது குறிப்பிடும்படி அமையாமல் (not significant) சோதனை முடிவுகளில் ஏற்பட்ட வித்தியாசம் மாதிரி எடுத்தலின் ஏற்றத் தாழ்வுகளுக்கானதா என்பதை ஈருறுப்புப் பரவலைப் பொருத்துதல் சோதனையில் (chi-square test-goodness of fit) காண உள்ளோம்.

ஈருறுப்புப் பரவலை உபயோகப்படுத்தும்போது, ஒப்பீட்டு அடிப்படையில் அமைந்த நிகழ்வெண்களை (relative frequency) அல்லது நிகழ்வெண் விகிதங்களைக் (frequency ratio) கையாளும் நிலை அடிக்கடி ஏற்படும். மேற்கண்ட சோதனையின் வெற்றிகளின் எண்ணிக்கைகளுக்கு மாறாக வெற்றி விகிதச் சமங்களாக 0/10, 1/10,...... எனவும் குறிப்பிடலாம். இம் முறைப்படி கணக்கிடும் போது, பிரிவு நிகழ்வெண் மாறுதமைந்து, பரவலின் கூட்டுச் சராசரி 'p' ஆகவும், திட்டவிலக்கம் $\sqrt{pq/n}$ ஆகவும் அமையும்.

புள்ளியியல் முறையில் நமக்குக் கிடைத்துள்ள வடிவங்களில் ஈருறுப்புப் பரவலும் ஒன்று. இதுபோன்று உருவாக்கப்பட்ட

வடிவங்கள் சில துறைகளில் நிகழும் இயற்கை நிகழ்ச்சிகளுக்குப் பொருந்துமாறு அமைவதைக் காணலாம். இம்மாதிரி இடங்களில் இவ் வடிவத்தின்மூலம் நிகழ்ச்சிகள்பற்றிப் புரிந்துகொள்ளவும், பொதுமையான முடிவுகள் எடுக்கவும் முடிகிறது.

காட்டு :

ராண்டமாக எடுக்கப்பட்ட 8 எலிகள் *A*, *B* என்று இரு சம பகுதிகளாகப் பிரிக்கப்பட்டுள்ளன. *A* என்ற பகுதியில் உள்ள 4 எலிகளுக்கு நான்கில் ஒன்றைக் கொல்லும் சக்தி வாய்ந்த நஞ்சும், *B* என்ற பகுதியில் உள்ள 4 எலிகளுக்கு நான்கில் இரண்டைக் கொல்லும் நஞ்சும் கொடுக்கப்படுகிறது. *B* பகுதியைவிட *A* பகுதியில் அதிகமான எலிகள் இறப்பதற்கான வழிகளையும் அவற்றின் நிகழ்திறன்களையும் கணக்கிடு.

B பகுதியில் *A* பகுதியைவிடக் குறைவான எலிகள் இறப்பதற்கான வழிகள்:

A-பகுதியில் இறக்கும் எலிகளின் எண்ணிக்கை	*B* பகுதியில் இறக்கும் எலிகளின் எண்ணிக்கை
0	பூச்சியத்திற்கு அதிகமாக (1, 2, 3, 4).
1	ஒன்றுக்கும் அதிகமாக (2, 3, 4).
2	இரண்டிற்கும் அதிகமாக (3, 4)
3	மூன்றிற்கும் அதிகமாக (4).

மேற்கண்ட முறைகளில் நிகழும் நிகழ்ச்சிகளுக்கான நிகழ்திறன்கள் *A*, *B* ஆகிய இரு பிரிவுகளுக்குமான இரு ஈருறுப்புக் கோவைகளைப் பயன்படுத்திக் கணக்கிடலாம்.

A-பிரிவு

ஓர் எலி இறப்பதற்கான நிகழ்திறன் $= p = \frac{1}{4}$.

எலி இறக்காமலிருப்பதற்கான நிகழ்திறன் $= q = 1 - p = \frac{3}{4}$.

$$(p + q)^n = \left(\frac{1}{4} + \frac{3}{4}\right)^4$$

$$= \frac{1}{256} + \frac{12}{256} + \frac{54}{256} + \frac{108}{256} + \frac{81}{256}$$

B-பிரிவு

ஓர் எலி இறப்பதற்கான நிகழ்திறன் $= p = \frac{1}{2}$.

எலி இறக்காமலிருப்பதற்கான நிகழ்திறன் $q = 1 - p = \frac{1}{2}$.

$$(p + q)^n = \left(\frac{1}{2} + \frac{1}{2}\right)^4$$

$$= \frac{1}{16} + \frac{4}{16} + \frac{6}{16} + \frac{4}{16} + \frac{1}{16}$$

B-பிரிவில் ஓர் எலியும் இறக்காமலிக்கும்போது, *A* பிரிவில் நிகழும் 1, 2, 3, 4 எலிகள் இறப்பதற்கான நிகழ்திறன்

$$= \left(\frac{1}{16} \times \frac{108}{256}\right) + \left(\frac{1}{16} \times \frac{54}{256}\right) + \left(\frac{1}{16} \times \frac{12}{256}\right) + \left(\frac{1}{16} \times \frac{1}{256}\right) = \frac{175}{4096}$$

B-பிரிவில் ஓர் எலியும் *A*-பிரிவில் 2, 3, 4 எலிகளும் இறப்பதற்கான நிகழ்திறன்

$$= \left(\frac{4}{16} \times \frac{54}{256}\right) + \left(\frac{4}{16} \times \frac{12}{256}\right) + \left(\frac{4}{16} \times \frac{1}{256}\right)$$

$$= \frac{268}{4096}$$

B-பிரிவில் இரண்டு எலிகளும் *A*-பிரிவில் 3, 4 எலிகளும் இறப்பதற்கான நிகழ்திறன்

$$= \left(\frac{6}{16} \times \frac{12}{256}\right) + \left(\frac{6}{16} \times \frac{1}{256}\right) = \frac{78}{4096}$$

B-பிரிவில் மூன்று எலிகளும் *A*-பிரிவில் 4 எலிகளும் இறப்பதற்கான நிகழ்திறன்

$$= \left(\frac{4}{16} \times \frac{1}{256}\right) = \frac{4}{4096}$$

ஒன்றையொன்று விலக்கும் தன்மைத்தான இந் நிகழ்ச்சிகளில் ஏதாவது ஒன்று நிகழ்வது *B* பிரிவைக் காட்டிலும் *A* பிரிவில் அதிகமான எலிகள் இறப்பதைக் காட்டும்.

$$\text{நிகழ்திறன்} = \frac{175}{4096} + \frac{286}{4096} + \frac{78}{4096} + \frac{4}{4096}$$

$$= \frac{525}{4096}$$

காட்டு :

பின்இணைப்பில் கொடுக்கப்பட்ட ராண்டம் எண்களில் 5 எண்கள் கொண்ட நூறு பகுதிகள் சோதனைக்குள்ளாக்கப்பட்டன. ஒவ்வொரு 5 எண்கள் கொண்ட பகுதியிலும் 0, 1 வருவது வெற்றியையும் மற்ற எண்கள் வருவது தோல்வியையும் குறிக்கட்டும்.

$n = 5, p = .2$ என அமைந்த 100 நிகழ்ச்சிகளுக்கான பரவலை அடியிற்காணும் பட்டியல் தருகிறது.

வெற்றி	நிகழ்வெண்
0	32
1	44
2	17
3	6
4	1
5	0
	100

இதற்கான கருதுகோள் நிகழ்வெண்களைக் கணக்கிட $p^x \; q^{n-x}$ கணக்கிட்டுக் கொள்கிறோம்.

$x = 0$ வெற்றியில்லாதது $= q^n = (0 \cdot 8)^5 = 0 \cdot 32768$

$x = 1$ $p^1 q^{n-1}$ $= (0 \cdot 2)(\cdot 8)^4 = 0 \cdot 08192$

$pq^{n-1} = \frac{p}{q} \; q^n$ என எழுதி, pq^{n-1}-ன் மதிப்பை q^n-ன் மதிப்பிலிருந்தும் கணக்கிட்டுக் கொள்ளலாம்.

$$\frac{p}{q} = \frac{\cdot 2}{\cdot 8} = \frac{1}{4}$$

$$pq^{n-1} = q^n \cdot \frac{p}{q} = \frac{0 \cdot 32768}{4} = 0 \cdot 08192$$

இதேபோல்

$x = 2$, $p^2 \; q^{n-2}$ என்பது $x = 1$ என்பதற்கான மதிப்பை $\frac{p}{q}$-வால் பெருக்குவதற்குச் சமமாகும். இந் நிகழ்ச்சி நிகழ்வதற்கான மொத்த வழிகள் (ஈருறுப்புக் குணகங்கள்) பாஸ்கல் முக்கோணத்தில் (Pascal's triangle) $n = 5$ என்பதற்கு நேராக அமைந்த மதிப்புகளுக்குச் சமமாகும். (1, 5, 10, 10, 5, 1) கருதுகோள் நிகழ்வெண் தரும் பட்டியல் பின்வருமாறு:

x	$p^x q^{n-x}$	$\binom{n}{x}$	$\binom{n}{x} p^x q^{n-x}$	$N\binom{n}{x} p^x q^{n-x}$	
0	0·32768	1	0·32768	32·77	32
1	0·08192	5	0·40960	40·96	44
2	0·02048	10	0·20480	20·48	17
3	0 00512	10	0·05120	5·12	6
4	0·00128	5	0·00640	0·64	1
5	0·00032	1	0.00032	0·03	0
				100·00	100

பாய்ஸான் பரவல் (Poisson Distribution)

ஒரு நிகழ்ச்சி எவ்வளவு முறை நிகழும், எவ்வளவு முறை நிகழாதிருக்கும் என்பதை அறியும் சூழலில் ஈருறுப்புப் பரவல் உபயோகப்படுத்தப்படுவதைக் கண்டோம். சில இடங்களில் ஒரு நிகழ்ச்சி எவ்வளவு முறை நிகழும் என்பதைத் தெரிந்து கொள்வதைப்போல் அந் நிகழ்ச்சி நிகழாததற்கான முறைகளைத் தெரிந்துகொள்ளும் அவசியமோ அல்லது தெரிந்துகொள்ள முயல்வது பொருளற்றதாகவோ அமைகிறது. பாதையின் ஓர் இடத்தில் எவ்வளவு பேர் கடந்து சென்றார்கள் (குறிப்பிட்ட நேரத்தில்) என்பதைத் தெரிந்துகொள்வதற்கு எவ்வளவு பேர் கடக்காமல் இருந்தனர் என்பதைத் தெரிய முயல்வது தேவையற்றதாகும். ஒரு கிலோ விதையில் எவ்வளவு பழுதானது என்பதைத் தெரிய, எவ்வளவு பழுதற்றது என எண்ணுவது பொருளற்றதாகும். தொடர்ந்து நிகழும் நிகழ்ச்சிகளில் இது போன்ற நிகழ்ச்சிகள் அரிதான நிகழ்ச்சிகள் (rare events) ஆகும். இந் நிகழ்ச்சிகள் அரிதாக நிகழ்வதால், மிகக் குறைந்த நிகழ் திறனைப் பெற்றிருக்கும். திறமையான தட்டெழுத்தாளர் (typist) அடித்த ஒரு பக்கத்தில் உள்ள தவறுகளின் எண்ணிக்கை, நல்ல விதைகளின் ஒரு கிலோ எடையில் உள்ள பழுதான விதைகளின் எண்ணிக்கை, அரிதான வியாதியினால் பாதிக்கப்பட்டு ஒரு மாவட்டத்தில் ஓர் ஆண்டில் இறந்தவர்கள் எண்ணிக்கை

ஆகியவற்றைப் பாய்ஸான் நியதிக்குக் கட்டுப்பட்டவையாகக் கொள்ளலாம். அரிதாக நிகழும் நிகழ்ச்சிகள் ஒழுங்கற்றுத் தற் செயல் நிகழ்ச்சிகளாக (hapazard) அமையாமல் அவையும் ஒரு நியதிக்குக் கட்டுப்படுவதையே பாய்ஸான் பரவல் காட்டுகிறது.

நிகழ்ச்சிகளின் எண்ணிக்கை (n) மிக அதிகமாக அமைந்தும், நிகழ்ச்சி நிகழ்வதற்கான நிகழ்திறன் மிகக் குறைந்தும் ($p \longrightarrow 0$), இவ் விரு அளவுகளின் பெருக்குத் தொகை வரையுற்றதாயும் (np: finite number) அமையும்போது, ஈருறுப்புப் பரவலிலிருந்து பெறப்படும் பரவலே பாய்ஸான் பரவலாகும். இதுவும் ஒரு தனித்த மாறிக்கான (discrete variable) பரவலாகும்.

ஈருறுப்புப் பரவலின் கூட்டுச் சராசரி np என்பதை அறிவோம். இதன் மதிப்பை இங்கு ஒரு வரையுற்ற எண்ணுகக் கொள்கிறோம். $np = m$ எனக் கொண்டால் $p = \frac{m}{n}$ என அமையும்.

'n' நிகழ்ச்சிகளில் 'x' வெற்றி தரும் ஈருறுப்புப் பரவல்

$$p(x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!\,(n-x)!} p^x q^{n-x}$$

$$= \frac{n!}{x!\,(n-x)!} \left(\frac{m}{n}\right)^x \left(1-\frac{m}{n}\right)^{n-x}$$

$$\left[p = \frac{m}{n};\ q = (1-p) = 1-\frac{m}{n}\right]$$

$$= \frac{n!}{n^x\,(n-x)!} \frac{m^x}{x!} \left(1-\frac{m}{n}\right)^{n-x}$$

கொடுக்கப்பட்ட நிபந்தனைகளை மனத்திற்கொண்டு கணக்கியல் முறைப்படி சுருக்கினால், பாய்ஸான் நியதிப்படி 'x' வெற்றிகள் கிடைப்பதற்கான நிகழ்திறன்

$$p(x) = \frac{m^x e^{-m}}{x!}.$$

அடுத்தடுத்து வரும் உறுப்புகளின் நிகழ்திறன், $e^{-m} \Sigma \frac{m^x}{x!}$ ஐ விரித்தால் கிடைக்கும் உறுப்புகளைக் குறிக்கும்.

$$P(x)=\frac{m^x}{x!}e^{-m}=e^{-m}\left(1+\frac{m}{1!}+\frac{m^2}{2!}+\ldots\ldots+\frac{m^x}{x!}+\ldots\ldots\right)$$

x, 0, 1, 2, ... என்ற மதிப்புகளைப் பெறுவதாயும், *m*ஐக் கூட்டுச் சராசரியாகவும் கொள்கிறோம். நிகழ்திற மதிப்புகளைக் கருதுகோள் நிகழ்வெண்களாக (Theoritical frequencies) மாற்ற N ஆல் பெருக்க வேண்டும்.

மாறி x	நிகழ்திறன் $p(x)$	கருதுகோள் நிகழ்வெண் $N \cdot p(x)$
0	e^{-m}	$N \cdot e^{-m}$
1	$m\,e^{-m}$	$N \cdot m \cdot e^{-m}$
2	$\frac{m^2}{2!}\,e^{-m}$	$N \cdot \frac{m^2}{x!}\,e^{-m}$
.	.	.
.	.	.
.	.	.
x	$\frac{m^x}{x!}\,e^{-m}$	$N\,\frac{m^x}{x!}\,e^{-m}$

பாய்ஸான் பரவலின் கூட்டுச் சராசரியும், திட்டவிலக்க வர்க்கமும் (variance) சமமாக அமைகின்றன. சோதனையில் கிடைக்கும் பரவல் ஒன்றின் கூட்டுச் சராசரியும், திட்டவிலக்கமும் சமமாக அமையின், இப் பரவலுக்கும் பாய்ஸான் பரவலுக்கும் உள்ள வேறுபாடு மிகக் குறைந்து இருக்கும்.

பாய்ஸான் பரவலின் கூட்டுச்சராசரி = திட்டவிலக்க வர்க்கம் = m.

ஒரு நிகழ்ச்சி நடைபெறுவதற்கான எண்ணிக்கை தெரியாமல் இருக்கும் இடங்களில், குறிப்பிட்ட நிபந்தனைகளுக்கொத்த சோதனைகளில் பாய்ஸான் பரவல் பெரிதும் உபயோகப்படும். இத்தகைய நிகழ்ச்சிகள் உயிரியல் ஆராய்ச்சிப் பிரிவில் பெரிதும் நிகழ்வதால் உயிர்ப் புள்ளியியலில் பாய்ஸான் பரவலின் உபயோகம் வலியுறுத்தப்படுகிறது.

காட்டு :

ஒரு கலவையில் உள்ள காடிச் சத்தணுக்களைச் (east cell) செறிவு குன்றியதாக்கி, நன்கு குலுக்கி எண்ணத்தக்க 400 அறைகளில் நிரப்பி, நுண்நோக்காடி (microscope) உதவியால் எண்ணிய தன் முடிவுகள் பட்டியலில் தரப்பட்டுள்ளன.

ஒரு சதுரத்தில் அமைந்த செல்கள்	0	1	2	3	4	5	6	7	8	9	10
நிகழ்வெண்	103	143	98	42	8	4	2	0	0	0	0

சோதனையின் முடிவுகள் செல்களே இல்லாத அறைகள் 103-ம், ஒரு செல் இருந்த அறைகள் 143-ம் ஆறு செல் இருந்த அறைகள் இரண்டும், 7, 8, 9, 10 செல்கள் உள்ள அறைகள் ஒன்றும் இல்லை என்பதைக் காட்டுகின்றன. இம் முடிவுகளைப் பாய்ஸான் நியதிக்குக் கட்டுப்பட்டதாகக் கொண்டு பாய்ஸான் பரவல் மூலம் கிடைக்கும் கருதுகோள் நிகழ்வெண்களைக் கணக்கிடுவோம்.

பாய்ஸான் பரவலில் உள்ள 'm' கூட்டுச் சராசரியின் மதிப்பிற்குச் சமமாவதால், இப் பரவலுக்கான கூட்டுச் சராசரியைக் கணக்கிட்டுக் கருதுகோள் நிகழ்வெண்களைக் கணக்கிடுகிறோம்.

x	0	1	2	3	4	5	6	7	8	9	10	மொத்தம்
f	103	143	98	42	8	4	2	—	—	—	—	400
$x_i f_i$	0	143	196	126	32	20	12	—	—	—	—	529

$$\bar{x} = m = \frac{\Sigma x_i f_i}{\Sigma f_i} = \frac{529}{400} = 1 \cdot 32.$$

$$\text{கருது நிகழ்வெண்} = N.\ \frac{e^{-m}\ m^{x}}{x!} = p(x).$$

$$p(x) = 400 \times \frac{e^{-1\cdot 32}\ (1\cdot 32)^{x}}{x!}$$

$$p(0) = 400 \times \frac{e^{-1.32}(1.32)^0}{0\,!} = 400 \times 0.2674 = 106.96$$

$$p(1) = 400 \times \frac{e^{-1.32}(1.32)^1}{1\,!} = 400 \times 0.2674 \times 1.32 = 141.19$$

$$p(2) = 400 \times \frac{e^{-1.32}(1.32)^2}{2\,!} = \frac{400 \times .2674 \times 1.32 \times 1.32}{1 \times 2}$$
$$= 93.1854$$

$$p(3) = 400 \times \frac{e^{-1.32}(1.32)^3}{3!} = \frac{400 \times .2764 \times 1.32 \times 1.32 \times 1.32}{1 \times 2 \times 3}$$
$$= 41.001$$

$$p(4) = 400 \times \frac{e^{-1.32}(1.32)^4}{4\,!} =$$

$$\frac{400 \times .2674 \times 1.32 \times 1.32 \times 1.32 \times 1.32}{1 \times 2 \times 3 \times 4} = 13.5303$$

$$p(5) = 400 \times \frac{e^{-1.32}(1.32)^5}{5\,!}$$

$$= \frac{400 \times .2764 \times 1.32 \times 1.32 \times 1.32 \times 1.32 \times 1.32}{1 \times 2 \times 3 \times 4 \times 5} = 3.5719$$

$$p(6) = \frac{400 \times e^{1.32}(1.32)^6}{6\,!} =$$

$$\frac{400 \times .2764 \times 1.32 \times 1.32 \times 1.32 \times 1.32 \times 1.32 \times 1.32}{1 \times 2 \times 3 \times 4 \times 5 \times 6} = 0.7858$$

$$p(7) = \frac{400 \times e^{-1.32}(1.32)^7}{7\,!} = 0.1493$$

$$p(8) = \frac{400 = e^{-1.32}(1.32)^8}{8\,!} = 0.0246$$

$$p(9) = 400 \times \frac{e^{-1.32}\,(1{\cdot}32)^{9}}{9\,!} = 0{\cdot}0037$$

$$p(10) = \frac{400 \times e^{-1.32}\,(1{\cdot}32)^{10}}{10\,!} = 0{\cdot}0005$$

ஒரு சதுரத்தில் அமைந்த நிகழ்வெண்	0	1	2	3	4	5	6	7	8	9	10	மொத்தம்
சோதனை நிகழ்வெண்	103	143	98	42	8	4	2	.	.	.	.	400
கருதுகோள் நிகழ்வெண்	107	141	93	41	14	4	1	.	.	.	.	400

கருதுகோள் நிகழ்வெண்கள் சோதனை நிகழ்வெண்களுடன் பாய்ஸான் நியதிக்குக் கட்டுப்பட்டுள்ளதா என்பதைக் கை-வர்க்கச் (chi-square test) சோதனைமுறைப் பகுதியில் காண உள்ளோம்.

காட்டு :

கால்நடைத் தீவனமாகப் பயன்படும் மணல் புல்வகை விதைகளில் (clover seed) ஒரு சதம் களைச் செடிகள் (weed seed) இருப்பதாகக் கணக்கிடப்பட்டுள்ளது. 100 விதைகள் கொண்ட மாதிரியில் 0, 1, 2, 3, ... களைச்செடிகள் இருப்பதற்கான நிகழ்திறன்களைக் கணக்கிடு.

களைச் செடிகள் இருப்பதற்கான தன்மையை ஒரு பாய்ஸான் நியதியாகக் கொண்டு 0, 1, 2......களைச்செடிகள் இருப்பதற்கான நிகழ்திறன்களைக் கணக்கிடுகிறோம். பாய்ஸான் பரவலுக்கான கூட்டுச்சராசரியைக் ($\bar{x} = np = m$) கணக்கிட்டு நிகழ்திற மதிப்புகள் கணக்கிடுகிறோம்.

கூட்டுச் சராசரி $= m = np = 100\,({\cdot}01) = 1.$

களைச்செடிகளின் எண்ணிக்கை	நிகழ்திறன்	
0	e^{-m}	0·3679
1	$m\ e^{-m}$	0·3679
2	$\frac{m \cdot m}{1 \cdot 2}\ e^{-m}$	0·1839
3	$\frac{m^3}{3!}\ e^{-m}$	0·0613
4	$\frac{m^4}{4!}\ e^{-m}$	0·0153
5	$\frac{m^5}{5!}\ e^{-m}$	0·0031
6	$\frac{m^6}{6!}\ e^{-m}$	0·0005
7	$\frac{m^7}{7!}\ e^{-m}$	0·0001

பயிற்சி :

1. செறிவு குன்றியதாக்கப்பட்ட ரத்தத்தில் எண்ணப்பட்ட வெள்ளை அணுக்கள் 121 என்ற கூட்டுச் சராசரியுடன் ஒரு பாய்ஸான் பரவலை ஒத்து அமைவதாகக் கொள்ளப்படுகிறது. 100-ம் அதற்குக் குறைவான வெள்ளை அணுக்கள் அமைவதற்கான நிகழ்திறனைக் கணக்கிடு.

2. ஒரு கன செ.மீ. அளவுள்ள ஒரு குறிப்பிட்ட திரவத்தில் 3 பாக்டீரியாக்கள் இருப்பதாகக் கணக்கிடப்பட்டுள்ளது. பத்து ஒருகன செ.மீ. சோதனைக் குழாய்களில் திரவம் நிரப்பப்பட்டுள்ளது. பாக்டீரியாக்கள் உண்டாகும் நிலை ஒரு பாய்ஸான் நியதிக்குக் கட்டுப்பட்டதெனக் கொண்டு (a) எல்லாச் சோதனைக் குழாய்களிலும் ஒரு பாக்டீரியாவாவது இருப்பதற்கான நிகழ்திறன், (b) ஏழு சோதனைக் குழாய்களில்தான் பாக்டீரியாக்கள் தோன்றியுள்ளன என்பதற்கான நிகழ்திறன்களைக் கணக்கிடு.

3. குதிரை உதைக்குப் பலியானவர்களின் நிகழ்வெண் பரவல் தரப்பட்டுள்ளது. புகழ்பெற்ற இக் கருத்துச் சேர்க்கை 20 ஆண்டுகளில் 10 பெர்ஸியன் தரைப்படையில் குதிரை உதைக்குப் பலியானவர்களைக் குறிக்கும். 200 ஆண்டுகளில் ஒரு தரைப்படையில் நிகழ்ந்த நிகழ்ச்சிகளாகக் குறிக்கப்பட்டுள்ளன. இது ஒரு பாய்ஸான் நியதிக்குக் கட்டுப்பட்டதாகக் கொண்டு கருதுகோள் நிகழ்வெண்களைக் கணக்கிடு.

x	f
0	109
1	65
2	22
3	3
	1
மொத்தம்	200

4. 400 அறைகளில் எண்ணப்பட்ட காடிச் சத்தணுக்களின் நிகழ்வெண் பரவல் தரப்பட்டுள்ளது. நிகழ்ச்சிகளின் தன்மை பாய்ஸான் நியதியைத் தழுவியதாகக் கொண்டு கணக்கியல் எதிர் பார்ப்பு நிகழ்வெண்களைக் கணக்கிடு.

காடிச் சத்தணுக்களின் எண்ணிக்கை	0	1	2	3	4	5	மொத்தம்
நிகழ்வெண்	213	128	37	18	3	1	400

5. நெற்பாளேகளேக் கணக்கிடச் செய்யப்பட்ட சோதனே முடிவுகள் தரப்பட்டுள்ளன. பாளேயின் நிகழ்வு, பாய்ஸான் நியதிக்குக் கட்டுப்பட்டதாகக் கொண்டு கொடுக்கப்பட்ட முடிவுகளுக்கான எதிர்பார்ப்பு நிகழ்வெண்களேக் கணக்கிடு.

மாதிரியில் உள்ள பாளேகளின் எண்ணிக்கை	மாதிரிகளின் எண்ணிக்கை
0	316
1	77
2	12
3	—
4	—
மொத்தம்	405

6. விதைப் பொட்டலமொன்றில் உள்ள விதைகளில் மூன்றில் ஒரு பகுதி சிவப்புநிற மலர்களேயும், மற்றவை வெண்மை நிற மலர்களேயும் கொடுக்கும் எனக் குறிக்கப்பட்டுள்ளது. ஏழு விதைகள் ஒரு பொட்டலத்திலிருந்து எடுக்கப்படுகின்றன. அவற்றிலிருந்து கிடைக்கும் மலர்கள்

(a) ஒன்றுமே சிவப்புநிற மில்லாதிருப்பது
(b) ஒன்றில்மட்டும் சிவப்பு நிறம் வருவது
(c) ஒன்றிலும் வெண்மை நிறம் இல்லாமல் இருப்பது ஆகியவற்றிற்கான நிகழ்திறன்களேக் கணக்கிடு.

7. 20 சதம் கறுப்பு மற்றும் 80 சதம் வெள்ளே அவரைகள் உள்ள ஒரு பை நம்மிடம் இருப்பதாகக் கொள்வோம். 8 அவரைகள் எடுக்கப்பட்ட ஒரு நிகழ்ச்சியில் ஆறும் அதற்கு அதிகமான வெள்ளே அவரைகள் அமைவதற்கான நிகழ்திறனேக் கணக்கிடு

8. ஆப்-டைப் (off-type) செடிகளைக் காணுவதற்கான சோதனையில் 10,000 விதைகள் விதைக்கப்பட்டன. முழுமைத் தொகுதியில் 1,000 செடிகளுக்கு 3 செடிகள் ஆப்-டைப்பாக இருப்பின், சோதனையிலும் 3 செடிகள் ஆப்-டைப்பாக இருப்பதற்கான நிகழ்திறனைக் கணக்கிடு.

9. 4 எலிகள் கொண்ட 103 ஒரீற்றுக் குட்டிகளில் உள்ள பெண் குட்டிகளுக்கான எண்ணிக்கை தரப்பட்டுள்ளது. தனித்த சோதனையில் பெண்குட்டி கிடைப்பதற்கான நிகழ்திறன் மாறிலியாக இருக்கும்போது, பெண் எலிகளுக்கான நிகழ்வெண் பரவலைக் கொண்டு கருதுகோள் நிகழ்வெண்களைக் கணக்கிடு.

பெண் எலிகளின் எண்ணிக்கை	நிகழ்வெண்
0	8
1	32
2	34
3	24
4	5
மொத்தம்	103

10. 3,500 விதைகளை ஏழு விதைகள்கொண்ட மாதிரிகளாகப் பிரித்துப் பயிர்செய்ததில் கிடைத்த முளைவிட்ட விதைகளின் நிகழ்வெண் பரவல் தரப்பட்டுள்ளது. ஈருறுப்புப் பரவலின் அடிப்படையில் கருதுகோள் நிகழ்வெண்களைக் கணக்கிடு.

முளைத்த விதைகளின் எண்ணிக்கை	0	1	2	3	4	5	6	7	மொத்தம்
நிகழ்வெண்	3	31	86	150	144	60	25	1	500

6. இயல்நிலைப் பரவல்

(Normal Distribution)

இயல்நிலை வளைகோட்டினால் அமைக்கப்படும் ஒரு மாறியின் அளவுகள், பொதுவாக மேலாராய்ச்சிக்கு அடிப்படையாக அமையும் ஊகக் கோட்பாடாகும். இயல்நிலை அளவுகள் மிகவும் அரிதாகக் கிடைக்கக்கூடிய தொன்றாகும். மற்ற துறைகளைக் காட்டிலும் உயிரியல் பிரிவில் கிடைக்கும் அளவுகள் இயல்நிலைப் பரவலுக்குப் பொருந்துவதாக அமைவதைக் காண முடியும். ஒத்த சூழலில் ஒருமைத்தான சத்துடைய உணவு உண்ணும் நபர்களின் உயரங்கள் (வயது மாற்றமிருப்பினும்) இயல்நிலைப் பரவலை ஒத்து அமைகின்றன. இயல் நிகழ்ச்சிகளில் இதுபோன்ற பரவல் அமைவதற்குக் காட்டாக ஒரே மர இலைகளின் நீளங்களின் பரவலைக் கூறலாம். சில மிக நீளமாகவும், சில குட்டையாகவும் அமைந்திருப்பினும், இலைகளின் நீளங்கள் யாவும் ஓர் எண்ணைச் சுற்றிக் குவிவதைக் காண முடியும்.

நியதிக்கணக்குகளை ஆராயும்போது, டி மோவ்ரேயால் (De Moivre) 1773ஆம் ஆண்டு இயல்நிலை வளைகோடு வரையறுக்கப்பட்டது. லாப்லாஸ் (Laplace, 1749—1827) மற்றும் காஸ் (Gauss, 1777—1855) என்ற இருவரும் இவ் வளைகோட்டினைத் தங்கள் துறை ஆராய்ச்சி முடிவுகளை விளக்கும்போது தனித் தனியே கண்டுபிடித்ததும் குறிப்பிடத்தக்கது. இவர்கள் அளவுகளில் ஏற்படும் பிழைகள் இயல்நிலைப் பரவலை ஒத்திருக்கிறது என அறுதியிட்டனர். இப் பரவல், 'இயல்நிலை நிகழ்திற வளைகோடு' (Normal probability curve), 'இயல்நிலைப் பிழைகளின் வளைகோடு' (Normal curve of error), இயல்நிலை மணி (∩) அமைப்பு வளைகோடு (Normal bell shapad curve), மற்றும் சமச்சீருடைய வளைகோடு (Symmetrical curve) எனப் பல பெயர்களால் குறிப்பிடப்படுகிறது. புள்ளியியல் பிரிவில் இயல்நிலைப் பரவல்

ஓர் இன்றியமையாத பரவலாக அமைந்து, கூறு எடுத்தல் நியதிக்கு அடிப்படைத் தத்துவமாகவும் அமைந்துள்ளது.

செயல் நிகழ்வதற்கான எண்ணிக்கை(n) அதிகப்படும்போதும், நிகழ்திறன் (p) மிகக் குறைவாக இல்லாதபோதும், ஈருறுப்புப் பரவலின் ஓர் எல்லை அமைப்பாக இயல்நிலைப் பரவல் அமைகிறது.

x என்ற ஈருறுப்பு மாறியிலிருந்து $z = \frac{x-np}{\sqrt{npq}}$ என்ற மாறியை அமைப்போம். np, $\sqrt{npq}$ ஈருறுப்புப் பரவலின் கூட்டுச் சராசரி மற்றும் திட்டவிலக்கத்தைக் குறிக்கும். z-மாறியின் கூட்டுச் சராசரி பூச்சியமாகவும், திட்டவிலக்கம் ஒன்றாகவும் அமையும். n-எண்ணிலி (infinity) ஆகும்போது z-மாறியின் பரவல் தொடர்ச்சிப் பரவலாக (continuous distribution) $(-\infty, +\infty)$ எல்லைக்குள் பூச்சியக் கூட்டுச் சராசரியையும் ஒன்று என்ற திட்டவிலக்கத்தையும் பெற்று அமையும்.

$$dp = \frac{1}{\sqrt{2\pi}}\, exp\left[-\tfrac{1}{2}\, z^2\right], \quad -\infty \le z \le \infty.$$

$$z = \frac{x-m}{\sigma}, \qquad \begin{array}{l} m : \text{கூட்டுச் சராசரி} \\ \sigma : \text{திட்ட விலக்கம்} \end{array}$$

என மாற்றி அமைக்கும்போது, இவ் வளைகோட்டின் அமைப்பு

$$dp = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\tfrac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right] \cdot dx$$

$$-\infty < x \le +\infty$$

$= f(x)\, dx$ என அமைகிறது.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, exp\left[-\tfrac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right]$$

$y=f(x)$ என்ற சமன்பாட்டிற்கான வளைகோடே இயல்நிலை நிகழ்திற வளைகோடு (Normal probability curve) எனக் கொள்ளப்படுகிறது. x எனும் தொடர்மாறி $y=f(x)$ எனும் நிகழ்திற அடர்த்திச் சார்பினை (probability density fn) 'm' என்ற கூட்டுச் சராசரியுடனும், 'σ' என்ற திட்டவிலக்கத்துடனும் பெறுகிறது. இதைக் கணக்கியல் குறியீடாக $x \sim N(m, \sigma)$ எனக் குறிக்கலாம்.

இயல்நிலைப் பரவலின் வளைகோடு அடியிற்காணும் பண்புகளைப் பெற்றிருப்பது குறிப்பிடத்தக்கதாகும். இதன் உச்சநிலை மதிப்பு (maximum ordinate) வளைகோட்டின் மத்தியில் அமைந்து

கூட்டுச் சராசரி, இடைநிலை ஆகியவையும் இப் புள்ளியில் மதிப்பைப் பெறுவதைக் காணலாம். சமச்சீரான இப் பரவலின் கூட்டுச்சராசரி இடைநிலை, முகட்டிற்கான மதிப்புகள் ஒரே புள்ளியில் அமையும்.

கூட்டுச் சராசரி, திட்டவிலக்கம் ஆகிய இரண்டளவைகளால் இயல்நிலை வளைகோடு முழுதும் வருணிக்கப்படுகிறது. கூட்டுச் சராசரியிலிருந்து திட்டவிலக்கத் தூரம் தள்ளி இரு புறங்களிலும் குறிக்கப்படும் புள்ளிகளிலிருந்து வரையும் நிலைக்கோடுகள் இயல் நிலை வளைகோட்டை வெட்டும் இடங்களே குவிவைக் குழிவாக மாற்றும் (points of inflexion) எனப்படும். இப் புள்ளிகளிலிருந்து வளைகோட்டுக்கு வரையும் தொடுகோடுகளின் சுழற்சிகள் இடமிருந்து வலமாக மாறுகின்றன.

கணக்கியலாக x-மாறியின் எல்லை $-\infty$ to $+\infty$ ஆக அமைவதாகக் கொண்டாலும் 0·03 சத அளவுகள்தான் $\bar{x} \pm 3\sigma$ எல்லையைக் கடந்த அளவுகளாக அமையும். அதாவது கொடுக்கப்பட்ட பரவலின் 99·97 அளவுகள் $x-3\sigma$, $x+3\sigma$ என்ற எல்லைக்குள் அமைவதை இது குறிக்கிறது.

இடைநிலையிலிருந்து கால்மானங்கள் சமதூரத்தில் அமைந்து கால்மான விலக்கம் 3-ல் 2 பங்கு திட்டவிலக்கத்திற்குச் சமமாக அமைவது இயல்நிலைப் பரவலின் மற்றொரு பண்பாகும்.

இயல்நிலை நிகழ்திறப் பரவலை, மூன்று முறையான மாறிகளால் குறிக்கலாம்.

x_i–கொடுக்கப்பட்ட அளவுகளை மாற்றமில்லாமல் கொள்வது.

$x_i - \bar{x}$– கொடுக்கப்பட்ட அளவைகளிலிருந்து கூட்டுச் சராசரியைக் கழித்துவரும் விலக்கங்களை மாறியாகக் கொள்வது.

$\frac{x_i - \bar{x}}{\sigma}$: கூட்டுச் சராசரியினின்றும் எடுக்கும் விலக்கங்களைத் திட்டவிலக்கத்தால் வகுத்துவரும் ஈவினை மாறியாகக் கொள்வது.

இறுதியாகக் குறிப்பிட்ட அளவுகளே $z_i = \frac{x_i - m}{\sigma}$ என்ற திட்ட மாறியாக இயல்நிலைப் பரவலில் உபயோகப்படுத்தப்படுகிறது. இவ் வளவுகள், கொடுக்கப்பட்ட அளவுகளின் அலகுகளிலிருந்து விடுபட்டு அமைவதால், இயல்நிலைப் பரவலுக்கான நிகழ்வெண்கள் மற்றும் பரப்புக்காணும் பட்டியலில் உபயோகப்படுகிறது.

$$dp = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right] dx \quad -\alpha \leq x \leq \alpha.$$

$z = \dfrac{x-m}{\sigma}$ எனக் கொண்டால்,

$$dp = \frac{N}{\sigma}\frac{1}{\sqrt{2\pi}} exp\left[-\frac{z^2}{2}\right] dz$$

$$= \frac{N}{\sigma}\phi(z)\, dz \quad -\alpha \leq z \leq a.$$

$$\phi(z) = \frac{1}{\sqrt{2\pi}} exp\left[-\frac{z^2}{2}\right]$$

$N=1,\ \sigma=1,\ m=0$ என அமையும்போது கிடைக்கும் இயல் நிலைப் பரவலே $dp = \phi(z)\, dz$ ஆகும். இந்த அமைப்பே இயல் நிலைப் பரவலின் திட்ட அமைப்பாக (standard form) கொள்ளப் படும். வளைகோட்டிற்குள் அமைந்த மொத்த இடத்தின் பரப் பளவு ஒன்றுதலின் இம் மாறியின் மதிப்பு z_1 மற்றும் z_2 என அமையும் இடத்திற்கிடையில் அமையும் இடத்தின் பரப்பு

$\int_{z_1}^{z_2}\phi(z)\, dz$ ஆகும். இதுவே z என்ற மாறி z_1 மற்றும் Z_2 என்ற எல்லைக்குள் அமையும் நிகழ்திறனாகும்.

இயல்நிலைப் பரவலின் நிகழ்வெண் பரப்புக் காண உதவும் பட்டியல் பின் இணைப்பில் கொடுக்கப்பட்டுள்ளது.

காட்டு :

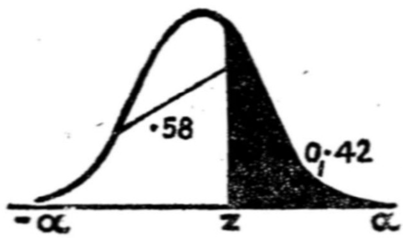
இயல்நிலைப் பரவலின் நியதியை அடிப்படையாகக் கொண்டு, மண்டையோடுகளின் நீள அகலக் குறியீட்டு எண்களில் 75-க்குக் குறைந்தவை *A* என்றும், 75—80 வரை உள்ளவை *B*-என்றும், 80-க்கு மேற்பட்டவை *C* என்றும் பிரிக்கப் பட்டுள்ளன.

A-58 சதமும், *B*-38 சதமும், *C*-4 சதமும் $f(z) = \dfrac{1}{\sqrt{2\pi}}\int_z^{\alpha} exp\left[-\frac{1}{2}t^2\right] dt$ எனவும் இருப்பதாகக்கொண்டு இப் பரவலுக் கான கூட்டுச் சராசரியையும் திட்டவிலக்கத்தையும் கணக்கிடு.

$z = 0{\cdot}2$ ஆக இருக்கும்போது, z-லிருந்து α வரை உள்ள இடத்தின் பரப்பு ·42 ஆகவும், z என்ற குத்துயரத்திற்கு இடப் புறம் அமைந்த இடத்தின் பரப்பு 0·58 ஆகவும் அமைகிறது. இந்தப் பரப்பு $x = 75$ என்பதைக் குறிக்கும்.

$$\frac{75-m}{\sigma} = Z$$

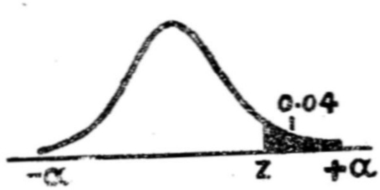
$$\frac{75-m}{\sigma} = 0{\cdot}2 \qquad (1)$$



படம் 6·1

$Z = \frac{80-m}{\sigma}$. குத்துயரத்திற்கு [$x = 80$ உள்ளபோது] வலப் புறம் அமைந்த இடத்தின் பரப்பு 0·04 எனக் கொடுக்கப்பட்டுள்ளது. குத்துயரத்திற்கு வலப்புறம் அமைந்த இடத்தின் பரப்பு ·04 என உள்ளபோது z-ன் மதிப்பு 1·75 எனக் கிடைக்கிறது.

$$\frac{80-m}{\sigma} = 1{\cdot}75 \qquad (2)$$



படம் 6.2

இரு சமன்பாடுகளிலிருந்தும் m, σ ஆகியவற்றின் மதிப்புகளைப் பெறலாம்.

$$\frac{75-m}{\sigma} = 0{\cdot}2 \qquad (1)$$

$$\frac{80-m}{\sigma} = 1.75 \qquad (2)$$

$$75-m = 0{\cdot}2\sigma$$

$$80-m = 1{\cdot}75\sigma$$

$$m+0{\cdot}2\sigma = 75 \qquad m = 74{\cdot}4$$

$$m+1{\cdot}75\sigma = 80 \qquad \sigma = 3{\cdot}2$$

கூட்டுச் சராசரி $= 74{\cdot}4$

திட்டவிலக்கம் $= 3{\cdot}2$

காட்டு :

ஒரு மாதிரியின் அளவுகள் $67{\cdot}56''$ கூட்டுச் சராசரியையும் $2{\cdot}75''$ திட்டவிலக்கத்தையும் கொண்டு ஓர் இயல்நிலைப் பரவலைத் தழுவி உள்ளன. $5'\ 2''$ உயரத்திற்குக் குறைந்த உயரம் உடைய வர்கள் எவ்வளவு பேர் என்பதைத் கணக்கிடு.

$$z = \frac{x-m}{\sigma} = \frac{62-67{\cdot}56}{2{\cdot}57} = -\ 2{\cdot}163$$

$z = 2{\cdot}163$ ஆக இருக்கும்போது நமக்குக் கிடைக்கும் பரப்பின் மதிப்பு

$$A = 0{\cdot}0153$$

$1{\cdot}5$ சதம்பேர் $62''$-க்குக் குறைவானஉயரம் உடையவர்களாக இருப்பார்கள்.

Z என்ற மாறி—1 முதல் + 1 வரை அமையும்போது, இயல்நிலைப் பரவலின் பரப்பைக் கணக்கிடுவோம்.

$$p(-1 < z \leqslant 1) = \int_{-1}^{+1} \phi(z)\, dz = 2\int_{0}^{1} \phi(z)\, dz.$$

$$= 2\times 0{\cdot}3413. \quad \left[\text{பட்டியல் மதிப்பு} \int_0^1 \phi(z)\, dz = 0{\cdot}3413\right]$$

$$= 0{\cdot}6826$$

அதாவது மொத்த நிகழ்ச்சிகளில் 68 சதத்திற்குமேல் $(-1, +1)$ இடைவெளிக்குள் அமையும்.

$$z = \frac{x-m}{\sigma} = -\ 1 \qquad x = m+\sigma.$$

$$z = \frac{x-m}{\sigma} = +\ 1 \qquad x = m-\sigma$$

மொத்த நிகழ்ச்சிகளில் 68 சத நிகழ்ச்சிகள் $(m-\sigma, m+\sigma)$ இடைவெளிக்குள் அமைவதை இது காட்டும்.

Z என்ற மாறி -2 முதல் $+2$ வரை அமையும்போது இயல்நிலைப் பரவலின் பரப்பு

$$P(-2 < z < +2) = \int_{-2}^{2} \phi(z)\, dz = 2\int_{0}^{2} \phi(z)\, dz.$$

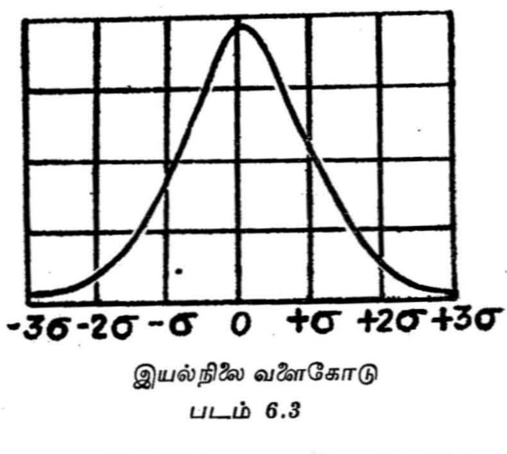
$$= 2 \times 0{\cdot}47725 = 0{\cdot}9545$$

மொத்த நிகழ்ச்சிகளில் 95% நிகழ்ச்சிகள் $(-2, +2)$ இடைவெளிக்குள் மதிப்புகளை ஏற்கின்றன. அதாவது மொத்த நிகழ்ச்சிகளில் 95 சத நிகழ்ச்சிகளின் மதிப்பு $(m-2\sigma;\ m+2\sigma)$ இடைவெளிக்குள் அமையும்.

Z என்ற மாறி -3 முதல் $+3$ வரை அமையும்போது இயல்நிலைப் பரவலின் பரப்பு

$$P(-3 \leq z \leq 3) = \int_{-3}^{3} \phi(z)\, dz = 2 \times 0{\cdot}4986 = 0{\cdot}9973.$$

அதாவது 99 சத நிகழ்ச்சிகள் $(-3, +3)$ பிரிவு இடைவெளிக்குள் மதிப்புகளைப் பெறும். எனவே, 99% நிகழ்ச்சிகள் $(\bar{x}-3\sigma,\ \bar{x}+3\sigma)$ பிரிவு இடைவெளியில் மதிப்புகளை ஏற்கின்றன.



இயல்நிலை வளைகோடு

படம் 6.3

$\bar{x} \pm \sigma$ இடைவெளியில் அமையும் பரப்பு 68%

$\bar{x} \pm 2\sigma$ இடைவெளியில் அமையும் பரப்பு 95%

$\bar{x} \pm 3\sigma$ இடைவெளியில் அமையும் பரப்பு 99%

நிகழ்திறப் பிழை (Probable error)

ஓர் உயிர்ப் புள்ளியியல் வல்லுநர் தாம் எடுக்கும் சோதனை களின் முடிவுகளைக் கொண்டு கூறும் கருத்துகட்கும் முடிவுகட்கும் நாம் கொடுக்கக் கூடிய நம்பிக்கையின் அளவைக் கூற முடியும் அளவைகளைக் கொடுப்பதே புள்ளியியலின் வேலையாகும். பெரும் பாலான இடங்களில் நமது நம்பிக்கைக் கூறுகளின் எண்ணிக்கை யைப் பொறுத்து அமைகிறது. கூறுகளின் எண்ணிக்கை மிகக் குறைவாக இருப்பின், கொடுக்கப்படும் முடிவுகள் கூறுகளின் ஏற்றத் தாழ்வுகளுக்கு உட்பட்டதாக அமைவதை உணர்வோம்.

ஒரு நாட்டின் மொத்த மக்கள் தொகையில், ஒரு நாளில் இறப்பவர்தம் சராசரி வயது, மற்றொரு நாளில் இறப்பவர்களின் சராசரி வயதினுக்குச் சமமாக அமையவேண்டுமென எதிர் பார்ப்பது நியாயமானதாகும். தவிர்க்கமுடியாத காரணங் களால் ஏற்படும் சில கொடிய வியாதிகள் குறிப்பிட்ட வயதினர் களையே தாக்கினாலன்றிக் கூறுகளின் ஏற்றத்தாழ்வுகளுக்கு மாறான வேறு காரணங்களை, சராசரிகளில் உள்ள வித்தியாசத் திற்குக் காரணமாகச் சொல்ல முடியாது. பத்து நாள்களுக்கான இறந்தவர் எண்ணிக்கையும், கூட்டுச் சராசரி வயதும் கொடுக்கப் பட்டுள்ளன.

நாள்	இறந்தவர் எண்ணிக்கை	கூட்டுச் சராசரி வயது (ஆண்டுகளில்)
ஜனவரி 13	31	30·16
,, 14	40	43 80
,, 15	27	40·59
,, 16	48	48·21
,, 17	32	48·34
,, 18	41	51·90
,, 19	39	46·28
,, 20	31	52·39
,, 1	39	51·62
,, 22	57	39·40

30-லிருந்து 60 பேர் கொண்ட கூறுகளில், கூறுகளுக்கான கூட்டுச் சராசரிகளின் ஏற்றத்தாழ்வு 22 ஆண்டுக்குமேல் அமைந்துள்ளது. இந்த ஏற்றத்தாழ்வுகள், கருத்துகளை முடிவாக்குவதற்கு முன் நிகழ்திறப் பிழைகளை (probable error) உணர்வது முக்கியம் என்பதை வலியுறுத்துகின்றன.

முடிவுகளை நம்புவதற்கான அளவை அல்லது கூறுகளின் ஏற்றத்தாழ்வுகளில் ஏற்படும் 'விலக்க அளவை'யாக நிகழ்திறப் பிழை குறிக்கப்படுகிறது. கூட்டுச் சராசரி அளவைக் குறிக்கும் புள்ளிக்கு இரு புறங்களிலும் சமதூரத்தில் இருக்கும் இரு புள்ளிகளின் இடையில் அமையும் இடத்தின் பரப்பு 50 சத உறுப்புகளைக் கொள்ளும்படியான மதிப்பையே நிகழ்திறப் பிழை (P.E) எனக் கொள்கிறோம்.

$$P(\bar{x} - P.E \leq x \leq \bar{x} + P.E) = P(-z_1 < t < z_1)$$
$$= 0.5$$

$$\int_0^{z_1} \phi(z)\, dz = 0.25$$

பட்டியலில் 0·25 பரப்பிற்கு ஏற்ற z-ன் மதிப்பு

$$z_1 = 0.6745.$$
$$= 2/3 \text{ தோராயமாக.}$$

$$z = \frac{x-m}{\sigma} = 2/3.$$

$$x - m = 2/3\, \sigma$$

நிகழ்திறப் பிழை $= | x_i - \bar{x} | = \frac{2}{3}\sigma.$

இதுவே கால்மான விலக்கத்திற்குச் சமமாக அமைவதைக் காணலாம். கால்மான விலக்கமும் $\left(Q.D = \frac{Q_3 - Q_1}{2}\right)$ கொடுக்கப்பட்ட நிகழ்ச்சிகளில் 50 சதம் உள்ளடங்கும் பரப்பைக் காட்டுவதாகும் என்பதை அறிவோம்.

கொடுக்கப்பட்ட இறப்புப் பட்டியலுக்கான கூட்டுச் சராசரி வயது 39·83 ± 2·60 எனக் கொள்ளின், இதுவே ஒரு கூட்டுச் சராசரி வயது 37·23 (39·83 − 2·60) மற்றும் 42·43 (39·83 + 2·90) இடைவெளிக்குள் அமையும் நிகழ்திறன், அக் கூட்டுச் சராசரியின் அளவு இவ் விடைவெளிக்குள் அமையாதிருப்பதற்குச்

சமமாகும். இரு அளவைகளுக்கு உள்ள வித்தியாசம் இருமடங்கு நிகழ்திறப் பிழைக்குக் குறைவாயின் வித்தியாசம் குறிப்பிடத் தக்கதாய் இல்லாமல், ஏற்பட்ட வித்தியாசம் கூறெடுத்தலின் ஏற்றத்தாழ்வுகளால் நிகழ்ந்தவை என்றும், மூன்று மடங்கு நிகழ் திறப் பிழைக்கு அதிகமாயின் வித்தியாசம் புறக்கணித்துவிடத் தக்க நிலையில் இல்லை எனவும் கொள்வோம்.

இயல்நிலைப் பரவலுக்கான கணக்கியல் நிகழ்வெண்களைக் காணல்

கொடுக்கப்பட்ட இயல்நிலைப் பரவலுக்குக் கருதுகோள் நிகழ் வெண்களைக் காண்பது இன்றியமையாததாய் உள்ளது. நிகழ் வெண்களைக் காணும் முறையைக் கீழ்க்கண்ட அட்டவணை விளக்கு கிறது.

Mid x	f	$x_i - x$	$z = \frac{x - x}{\sigma}$	$\phi(z)$	$y = \frac{N}{\sigma}\phi(z)$	$y \times c$	கருது கோள் நிகழ் வெண்	சோதனை முடிவு கள்

காட்டு :

பிரி வெண்	61	62	63	64	65	66	67	68	69	70	71	72	73	74
நிகழ் வெண்	2	10	11	38	75	93	106	126	109	87	75	28	9	4

கொடுக்கப்பட்ட நிகழ்வெண் பரவலுக்குக் கூட்டுச் சராசரி, திட்டவிலக்கம் கணக்கிடப்பட்டு, கருதுகோள் நிகழ்வெண் பரவலுக்கான பட்டியல் அமைக்கவேண்டும்.

x பிரிவெண்	f_i நிகழ் வெண்	$x_i f_i$	$x_i f_i^2$
61	2	122	7442
62	10	620	38440
63	11	693	43659
64	38	2432	155648
65	57	3705	240825
66	93	6138	405108
67	106	7102	475834
68	126	8568	582624
69	109	7521	518949
70	87	6090	426300
71	75	5325	378075
72	23	1656	119232
73	9	657	47961
74	4	296	21904
	750	50925	3462001

$$\bar{x} = \frac{50925}{750} = 67{\cdot}9$$

$$\sigma x^2 = \frac{3462001}{750} - (67{\cdot}9)^2$$

$$\sigma x = 2{\cdot}364$$

பிரிவெண்களை x-மாறியாகக் கொண்டு, மாறிகளினின்றும்' சராசரியைக் கழித்துவரும் மதிப்பை $(x_i - \bar{x})$ இரண்டாவதாக அமைக்கவேண்டும். காட்டாக x மாறி 66 என்ற பிரிவெண்ணைப் பெறும்போது $(x_i - \bar{x}) = -1 \cdot 9$ என அமையும்.

$\frac{(x_i - \bar{x})}{\sigma}$ என்ற அலகுகள் நீக்கப்பட்ட அளவுகளை அமைக்க வேண்டியது இதற்கடுத்த படியாகும்.

அலகுகள் நீக்கப்பட்ட அளவைகளுக்குச் சமமான மதிப்புகளை உடைய இயல்நிலைப் பரவல் மதிப்புகளை அட்டவணையைப் (Table of normal ordinate) பார்த்து அமைக்கவேண்டும். இம் மதிப்பை $\frac{N}{\sigma}$ c என்ற மதிப்பால் பெருக்கிவரும் தொகையே நாம் எதிர்பார்க்கும் நிகழ்வெண்ணாகும்.

பிரி வெண் x_i	$x - \bar{x}_i$	$\frac{x_i - \bar{x}}{\sigma} = z_i$	$\phi(z)$	$\frac{N}{\sigma} e \cdot \phi\ (z)$	கருது கோள் நிகழ் வெண்	சோதனை முடிவு
61	—6·9	—2·9188	·00568	1·7829	1·78	2
62	—5·9	—2·4958	·01795	5·6154	5·62	10
63	—4·9	—2·0728	·04660	14·7843	14·78	11
64	—3·9	—1·6497	·10230	32·4556	32·46	38
65	—2·9	—1·2267	·18810	62·8490	62·85	57
66	—1·9	—0·8037	·28910	91·6244	91·62	93
67	—0·9	—0·3807	·37110	117·7348	117·73	106
68	+0·1	+0·0423	·39855	126·4436	126·44	126
69	+1·1	+0·4653	·35810	113·6104	113·61	107
70	+2·1	+0·8883	·26898	85·3427	85·34	87
71	+3·1	+1·3113	·16890	53·5850	53·59	75
72	+4·1	+1·7343	·08870	28·1409	28·14	23
73	+5·1	+2·1574	·03890	12·4366	12·44	9
74	+6·1	+2·5804	·01430	4·5368	4·54	4

நிகழ்ச்சிகளின் எண்ணிக்கை மிக அதிகமாக அமையும்போது, இயல்நிலைப் பரவல், ஈருறுப்புப் பரவலுக்கு ஏற்றதாய் அமைவதைக் குறிப்பிட்டோம். $n = 30$ ஆக அமையும்போதே ஈருறுப்புப் பரவலுக்கும் இயல்நிலைப் பரவலுக்கும் வரையப்படும் நிகழ்வெண் செவ்வகங்கள் ஒத்து அமைகின்றன.

இயல்நிலைப் பரவலை ஒத்த முடிவுகள் நடைமுறையில் மிக அரிதாகக் கிடைப்பதால் இதன் முக்கியத்துவம் புள்ளியியல் பிரிவில் குறைந்து காணப்படும் என்று எண்ணினால் தவறான முடிவைக் கொண்டவர்களாவோம். மாறாக, இதன் முக்கியத்துவம் மிகவும் அதிகமாக உணரப்பட்டு வருவதையே காண முடிகிறது. பிழைகளின் நியதிப் பரவல்களும், உயிர்ப் புள்ளியியல் அளவுகளும் இயல்நிலைப் பரவலைப் பெரிதும் ஒத்து அமைவதைக் குறிப்பிட வேண்டியவர்களாக உள்ளோம்.

முழுமைத் தொகுதிக்கான பரவலின் தன்மையை எல்லா இடங்களிலும் நம்மால் அறிந்துகொள்ள முடிவதில்லை. இம்மாதிரி இடங்களில் இயல்நிலைப் பரவலாகக்கொண்டு எடுக்கப்படும் முடிவுகள் பெரும்பாலும் தவறின்றி அமைந்துள்ளன. தனித்த அளவுகளுக்கான பரவல்கள் இயல்நிலைப் பரவலை ஒத்து அமையாத இடத்தும், இவற்றிலிருந்து பெறப்படும் அளவைகள் ஓர் இயல்நிலைப் பரவலை ஒத்து அமைவது குறிப்பிடத்தக்கது. மாதிரிகளின் பரவல்கள் பெரும்பாலும் இயல்நிலைப் பரவலை ஒத்து அமைவதைக் காண உள்ளோம்,

கொடுக்கப்பட்ட மாறியின் அளவைகள் இயல்நிலைப் பரவலாக இல்லாத இடங்களில்கூட, மாறிகளை மாற்றியமைப்பதால் இயல்நிலைப் பரவல் கிடைக்கப்பெறுவது குறிப்பிடத்தக்கதாகும். x மாறி இயல்நிலைப் பரவலாய் இல்லாவிடினும், $\sqrt{x}$ இயல்நிலையைப் பெற்று அமையலாம். இத்தகைய பண்புகள் இதன் உபயோகத்தை வலியுறுத்துவதாய் அமைந்துள்ளன.

காட்டு :

கூட்டுச் சராசரி 30, திட்டவிலக்கம் 5 என உள்ள இயல்நிலைப் பரவலில் x-ன் மதிப்பு 20 மற்றும் 35 என உள்ள இடைவெளியில் நிகழும் உறுப்புகளின் எண்ணிக்கையைச் சதவீதத்தில் குறிப்பிடு.

$$x = 20. \quad \frac{x-m}{\sigma} = \frac{20-30}{5} = -\ 2{\cdot}0$$

$$x = 35. \quad \frac{x-m}{\sigma} = \frac{35\ -30}{5} = +\ 1{\cdot}0$$

இயல்நிலைப் பரவலின் பரப்புப் பட்டியலிலிருந்து (Normal table-area) $\int_{z} \phi(t)\,dt$, (− 2·00, 0·00) எனும் இடைவெளியில் அமைந்துள்ள பரப்பு 0·477250 எனவும், (0·00, 1·00) எனும் இடைவெளியில் அமையும் பரப்பு 0·34134 (·50000 = −0·15866) எனவும் பெற்று, (− 2·0, 1·0) எனும் இடைவெளியில் அமையும் பரப்பை 0·47725 + 0·34134 = 0·81859 எனக் காண்கிறோம்.

$$-\infty < z < -2 = 0{\cdot}0228.$$
$$-\infty < z < -1 = 0{\cdot}8185.$$
$$-2 < z < -1 = (-\infty < z < -2) - (-\infty < z < -1)$$
$$= 0{\cdot}8185 - 0{\cdot}228 = 20{\cdot}8185.$$

x = 20, x = 35 என்ற இடைவெளியில் அமையும் உறுப்புகளின் எண்ணிக்கை } = 81·85%

பயிற்சி :

1. கொடுக்கப்பட்ட அளவுகள் ஓர் இயல்நிலைப் பரவலைத் தழுவியதாகக் கொண்டு கருதுகோள் நிகழ்வெண்களைக் கணக்கிடு.

உயரம் செ.மீ.	நிகழ்வெண்
100–105	4
105–110	14
110–115	860
115–120	138
120–125	206
125–130	298
130–135	380
135–140	450
140–145	500
145–150	430
150–155	260
155–160	128
160–165	66
165–170	28
170–175	12
மொத்தம்	2974

2. ஒரு மரத்திலுள்ள 900 இலைகளின் நீளங்களை அளந்ததில் கூட்டுச் சராசரி 3 செ.மீ. எனவும், திட்டவிலக்கம் 0·5 செ.மீ. எனவும் கிடைக்கிறது. இதன் பரவலை ஓர் இயல்நிலையாகக் கொண்டு (1) 4.2 செ.மீ-க்கு அதிகமான நீளமுடைய இலைகள் எவ்வளவு என்றும், (2) 2 செ.மீ.க்குக் குறைவான நீளமுடைய இலைகள் எவ்வளவு எனவும் கணக்கிடு.

3. 800 முட்டைக்கோசு விதைகளின் முளைத்தன்மை அடியிற் காணும் பட்டியலில் தரப்பட்டுள்ளது. 8 விதைகள் கொண்ட நூறு வரிசைகளில் முளைத்த விதைகளின் நிகழ்வெண் பரவல் குறிக்கப்பட்டுள்ளது. நிகழ்ச்சியை ஓர் இயல்நிலை நியதியாகக் கொண்டு கருதுகோள் நிகழ்வெண்களைக் கணக்கிடு.

ஒரு வரிசையில் முளைத்த விதைகளின் எண்ணிக்கை	வரிசைகளின் எண்ணிக்கை
0	13
1	23
2	34
3	15
4	11
5	4
—	100

4. நூறு அளவுகள் 534 செ.மீ. கூட்டுச் சராசரியையும், 13·5 செ.மீ. திட்டவிலக்கத்தையும் கொண்டு ஓர் இயல்நிலைப் பரவலைத் தழுவியுள்ளதாகக் குறிக்கப்பட்டுள்ளன. 561 செ.மீ. அளவிற்கும் அதிகமான அளவுகளின் எண்ணிக்கை யாது? 520·5 செ.மீ., 547·5 செ.மீ. இடைவெளிக்குள் அமையும் அளவுகள், மத்தியில் அமையும் 50 சத உறுப்புகளுக்கான எல்லைகள் ஆகியவற்றைக் கணக்கிடு.

5. கூட்டுச் சராசரி (12), திட்டவிலக்கம் (3) என உள்ள ஓர் இயல்நிலைப் பரவல், கூட்டுச் சராசரியிலிருந்து ஒரு குறிப்பிட்ட அளவிற்கான நிகழ்திறன் 0·4115 இருக்கும் அளவின் மதிப்பைக் குறிப்பிடு. x-ன் மதிப்பு 3·5 என இருக்குமிடத்திலிருந்து, ஒரு குறிப்பிட்ட மதிப்பிற்கான அளவின் நிகழ்திறன் 0·2307 என அமைந்த அளவினைக் குறிப்பிடு.

6. கொடுக்கப்பட்ட நிகழ்வெண் பரவலுக்கான கருதுகோள் நிகழ்வெண்களைக் கணக்கிட்டு, $\bar{x} \pm \sigma$, $\bar{x} \pm 2\sigma$, $\bar{x} \pm 3\sigma$ எல்லைக்குள் அமையும் நிகழ்வெண்களின் சதவீதத்தைக் குறிப்பிடு.

பிரிவு இடைவெளி	நிகழ்வெண்
5·5–6·5	3
6·5–7·5	6
7·5–8·5	9
8·5–9·5	13
9·5–10·5	8
10·5–11·5	5
11·5–12·5	4
மொத்தம்	48

7. பண்புகளின் இயைபு

(Assosiation of Attributes)

புள்ளியியல் ஆய்விற்கான அடிப்படைச் செய்திகள் இரு வகைகளாகப் பிரிக்கப்படலாம் என்பதை முன்பு சுட்டினேம். நம் நாட்டின் மக்கள் தொகையை மதத்தின் அடிப்படையில் பிரிப்பது பண்பினை அடிப்படையாகக் கொண்ட பகுப்பாகவும், வயதின் அடிப்படையில் பிரிப்பது அளவினை அடிப்படையாகக் கொண்ட பகுப்பாகவும் கொள்ளப்படும். அளவுகளின் அடிப்படையில் சேகரிக்கப்பட்ட செய்திகளைப் பகுக்கும் முறைகளை இதுவரை கண்டோம். பண்பின் அடிப்படையில் சேகரிக்கப்படும் செய்திகளுக்கான தன்மையையும், அவற்றை ஆராயும் முறைகளையும் இங்குக் காண உள்ளோம்.

முழுமைத் தொகுதியிலுள்ள உறுப்புகளில் குறிப்பிட்ட பண்பு அமையப்பெற்றிருக்கிறதா அமையப்பெறவில்லையா என்பதைக் குறிப்பதே பண்புகளின் தன்மையைக் காணும் முதற்படியாகும். மக்கள் கணக்கெடுப்பில் ஆண், பெண், பார்வை உள்ளவர்கள், பார்வையற்றவர்கள் எனவும், நாணயத்தைச் சுண்டுவதில் கிடைக்கும் தலை, பூக்களின் எண்ணிக்கையையும், விதைப்பொட்டலங்களில் அமைந்த பழுதான, பழுதற்ற விதைகளின் எண்ணிக்கையையும் குறிப்பிடுவது பண்பின் அடிப்படையில் பிரிப்பதற்குக் காட்டாகச் சொல்லலாம். குறிப்பிட்ட பண்பு அமையப்பெற்ற உறுப்புகளின் எண்ணிக்கையைக் குறிப்பிடுவதிலேயே அவற்றுக்கான தன்மை அமையப்பெறுகிறது.

பண்பின் அடிப்படையில் முழுமைத்தொகுதி இரு பிரிவுகளாகப் பிரிக்கப்படுவது, பண்பிற்கான பகுப்பின் முதல் நிலையாகும். குறிப்பிட்ட பண்பு பெற்றவை ஒரு பிரிவாகவும், பெருதவை மற்ருெரு பிரிவாகவும் குறிக்கப்படும். ஒரு கல்லூரியில் உள்ள ஆண்கள் நமது கவனத்துக்குரிய பண்பாயின்,

1200 மாணவர்கள் உள்ள கல்லூரியில் 900 ஆண்கள் நமது பண்பைப் பெற்றவர்களாகவும் 300 பெண்கள் நமது பண்பைப் பெறாதவர்களாகவும் குறிக்கப்படுவார்கள். பண்பின் அடிப்படையில் பிரிக்கப்பட்ட இரு பிரிவுகளும், மேலும் பல பிரிவுகளாகப் பிரிக்கப்படலாம். கல்லூரியில் உள்ள 900 ஆண்களை வீட்டிலிருந்து வருவோர், விடுதியில் தங்கிப் படிப்போர் என மேலும் இரு பிரிவுகளாகவும் பிரித்துக் கொள்ளலாம். நமது தேவைக்கேற்பக் கொடுக்கப்பட்ட முழுமைத் தொகுதியைப் பண்பின் அடிப்படையில் எத்தனை பிரிவுகளாகவும் பகுத்துக் கொள்ளலாம்.

பண்பின் அடிப்படையில் பிரிக்கப் பின்வரும் கட்டளை விதிகள் இன்றியமையாதவையாக உள்ளன. (1) ஒரே தன்மைத்தான உறுப்புகளின் மொத்த எண்ணிக்கை தெரிந்திருக்கவேண்டும். (2) பகுக்க உதவும் பண்பு தெளிவானதாக நிகழ்திறப் பண்பு விளக்கத்தைப் பெற்றதாய் இருக்கவேண்டும். (3) பொருளைப் பார்ப்பதாலோ குறிப்புகளைப் படிப்பதாலோ பண்பு பெறப்பட்ட தன்மையும், பெறப்படாத தன்மையும் கணக்கிட முடிவதாக இருக்க வேண்டும். (4) பண்பு அமையப்பெற்ற உறுப்புகள் எண்ணக் கூடியனவாக இருக்கவேண்டும்.

முழுமைத் தொகுதியில் ஒரு குறிப்பிட்ட பண்பிற்கான தன்மையை அப் பண்பு அமையப்பெற்ற உறுப்புகளின் எண்ணிக்கையால் குறிக்கிறோம். 1951 மக்கள் கணிப்புப்படி தமிழ்நாட்டில் 10730 குடும்பங்களில் 6670 குடும்பங்கள் விவசாயம் செய்பவர்கள் எனக் குறிப்பிடப்பட்டுள்ளது. விவசாயம் செய்தல் நமது கவனத்திற்குரிய பண்பாயின் அப் பண்பின் தன்மை 6670 குடும்பங்களால் குறிக்கப்படும். விவசாயத் தொழில் பாகுபாடும் அதனைச் சார்ந்த மொத்த நபர்களும் காலத்தால் மாறும் தன்மைத்தன. மாறுபட்ட சூழலில் அமையும் இரு முழுமைத் தொகுதிகளை ஒப்பிடும்போது காலத்தாலும் இடத்தாலும் மாறும் தன்மைத்தான பண்புகளின் மொத்த உறுப்புகளை அவற்றின் விகிதச் சமங்களால் குறிப்பிடுவது நன்மை பயப்பதாகும். p என்பது பண்பு நிகழ்வதற்கான விகிதச் சமமாயின் $p = n_i/n$ என அமைந்து 'n' முழுமைத் தொகுதியின் மொத்த உறுப்புகளையும், n_i என்பது குறிப்பிட்ட பண்பு அமையப்பெற்ற உறுப்புகளையும் குறிக்கும்.

முழுமைத் தொகுதியிலுள்ள ஒவ்வோர் உறுப்பும் பண்பிற்குத் தக்கவாறு ஒன்று, இரண்டு அல்லது ஒன்றையொன்று விலக்கும் தன்மைத்தான பல பிரிவுகளில் ஏதாவதொன்றில்தான் அமையும்.

பிரச்சினைக்குரிய பொருளைப் பண்புகளின் அடிப்படையில் பிரிக்கும் போது இவ்வாறு அமைவதற்கேற்றபடி தெளிவுபடுத்திக் கொள்ளவேண்டும். ஆண் பெண், படித்தவர் படிக்காதவர் எனும் பாகுபாட்டில் சிக்கலில்லாமல் இருந்தாலும், வேலை யில்லாமை நமது சிக்கலுக்குரிய பண்பாயின், பாகுபாடு செய்யும் போது ஏற்படும் சிக்கல்களைத் தெளிவுபடுத்திக்கொள்ள வேண்டும். எத்தகைய நபர்கள் நமது பண்பைப் பெற்றவர்கள், எத்தகைய வர் நமது பண்பைப் பெறுதவர்கள் என்பதைக் கணக்கிடத் தெளிவான முறையை வகுத்துக்கொள்ள வேண்டும். கொடுக்கப் பட்ட முழுமைத் தொகுதியை யார் பாகுபாடு செய்யினும் ஒரே முடிவு அமையவேண்டிய நிலையில் நமது பண்பு தெளிவு படுத்தப்படுவது அவசியமானதாகும்.

சில இடங்களில் ஆண் பெண், படித்தவர் படிக்காதவர் என இரு பிரிவுகளாகப் பிரிக்கப்படுவதைக் காட்டிலும் தரப்படுத்திப் (rank) பிரிக்கப்படுவதும் உண்டு. ஒரு குறிப்பிட்ட முழுமைத் தொகுதியைச் செவிடர்கள், கேட்கும் தன்மையுடையவர்கள் என இரு கூருகப் பிரிப்பதைக்காட்டிலும் கேட்கும் தன்மையைத் தக்க கருவிகொண்டு அளந்து பல பிரிவுகளாகவும் பிரித்துக்கொள்ள லாம். இத் தன்மை பண்புகளின் அடிப்படையில் பிரிக்கும்போதும் இடத்திற்கேற்றுற்போல் அளவுகளின் குணத்தைப் பெறுவதைக் காட்டுகிறது.

பண்பின் அடிப்படையில் பிரிக்கும்போது இருவகையான கருதுகோள் சோதனைகள் (Testing of Hypotheses) செய்வதை மனத்தில் கொள்கிறோம். நமது ஆய்விற்குரிய பண்பு நம்மால் பிரிக்கப்படும் இரு மாதிரிகளிலும் குறிப்பிடத்தக்க வித்தியாசத் துடன் அமைந்துள்ளதா என்பதைக் காண்பது முதற் கருது கோட் சோதனையாகும். இரு நகரங்களில் உள்ள மொத்த மாணவர்களின் ஆண்/பெண் பாகுபாடு குறிப்பிடப்பட்ட வித்தியாசத்துடன் அமைந்துள்ளதா எனச் சோதனை செய்வதை இதற்குச் சான்றுகச் சொல்லலாம்.

இரு மாதிரிகளில் அமைந்த ஒரு பண்பிற்கும், ஒரே மாதிரி யில் அமைந்த இரு பண்புகளுக்கும் உள்ள இயைபைச் (assosiation) சோதனை செய்வது பண்புகளின் பாகுபாட்டில் செய்யப் படும் கருதுகோட் சோதனையாகும். ஒரு கிராமத்தில் பெரிய அம்மை தடுப்பூசி போட்டவர்களில் சிலருக்கு அம்மை கண்டது குறிப்பிடப்பட்டுள்ளது. இதுபோலவே தடுப்பூசி போடாத சிலருக்கும் அம்மை கண்டது குறிக்கப்பட்டுள்ளது. பண்பின்

அடிப்படையில் பிரிக்கப்பட்ட கிராம மக்கள்தம் தொகையிலிருந்து தடுப்பூசி போடுவதற்கும் அம்மை வராமலிருப்பதற்கும் உள்ள இயைபைக் காண்பதே நமது நோக்கமாகும். அம்மை வராமல் தடுப்பதில் தடுப்பூசிக்குள்ள சக்தியைப் பொறுத்தே அதை மக்கள் ஏற்றுக்கொள்ளும் நிலை உறுதிப்படுத்தப்படுவதால், இத்தகைய சோதனை ஓர் அடிப்படைத் தேவையாய் அமைகிறது.

இயைபைப் படிப்பதற்கும் கணிப்பதற்கும் உதவும் குறியீடுகளையும் அடிப்படைத் தேற்றங்களையும் காண்போம்.

ஒரு பண்பானது நேரிடைப் பண்பாகவோ (positive attribute), எதிரிடைப் பண்பாகவோ (negative attribute) இருக்கலாம். நாம் குறிப்பிடும் பண்பு அமையப்பெறுமாயின் அது நேரிடைப் பண்பு என்றும், இதற்கு மாறான பண்பு அமையப்பெறும் நிலை எதிரிடைப் பண்பு என்றும் சொல்லப்படும். முழுமைத் தொகுதியில் நம் ஆய்விற்குரிய பண்பினைப்பெற்ற உறுப்புகள் நேரிடைப் பண்புடையவை எனவும், அப் பண்பு அமையப் பெறாத உறுப்புகள் எதிரிடைப் பண்புடையவை எனவும் குறிக்கப்படும். பண்புகளை ஆராயும்போது நேரிடைப் பண்புகள் யாவும் A, B, C...... என்ற ஆங்கில எழுத்துகளாலும், எதிரிடைப் பண்புகள் யாவும் α, β, γ... என்ற கிரேக்க எழுத்துகளாலும் குறிக்கப்படும். A என்பது ஆண்களைக் குறித்தால் α என்பது பெண்களையும், B என்பது வீட்டிலிருந்து கல்லூரிக்கு வருவோரைக் குறித்தால் β என்பது மாணவர் விடுதியிலிருந்து கல்லூரிக்கு வருவோரையும் குறிக்கும். ஒன்றுக்கு மேற்பட்ட பண்புகள் அமையப்பெற்ற நிலையை, அவற்றுக்குச் சரியான குறியீடுகளைச் சேர்த்தெழுதிக் குறிப்பிடுகிறோம். $A\beta$ என்பது மாணவர் விடுதியிலிருந்து கல்லூரிக்குச் செல்லும் ஆண்களையும், αB என்பது வீட்டிலிருந்து கல்லூரிக்குச் செல்லும் பெண்களையும் குறிக்கும். முழுமைத்தொகுதி பிரிக்கப்பட்ட பண்புகள் அனைத்தும் எழுத்துகளால் சேர்த்தெழுதிக் குறிக்கப்படும். A என்பது பண்பினையும், (A) என்பது [அடைப்பிற்குள் அமையும் A] அப் பண்பமையப் பெற்ற நபர்களையும் குறிக்கும். எனவே, (A) என்பது A-பண்பு அமையப்பெற்ற உறுப்புகளின் நிகழ்வெண்ணைச் சுட்டும். (AB) என்பது A, B ஆகிய இரு பண்புகள் சேர்ந்தாற்போல் அமையப் பெற்றவர்கள் எண்ணிக்கையையும், $(\alpha\beta\gamma)$ என்பது A, B, C ஆகிய மூன்று பண்புகளும் சேர்ந்தாற்போல் அமையப்பெறாத உறுப்புகளின் எண்ணிக்கையையும் குறிக்கும்.

ஒரு பண்பு அமையப்பெற்றதைக் குறிக்கும் நிகழ்வெண் முதற்படி நிகழ்வெண் (First order frequency) எனவும், இரண்டு

பண்புகள் அமையப்பெற்றதைக் குறிக்கும் நிகழ்வெண்கள் இரண்டாம்படி நிகழ்வெண் (Second order frequency) எனவும், மூன்று பண்புகளைக் குறிக்கும் நிகழ்வெண் மூன்றும்படி நிகழ்வெண் (Third order frequency) எனவும் சொல்லப்படும். (A), (β) என்பன முதற்படி நிகழ்வெண்களாகவும் $(A\beta)$, (αB) என்பன இரண்டாம்படி நிகழ்வெண்களாகவும், $(A\beta C)$, $(\alpha B\gamma)$ என்பன மூன்றும்படி நிகழ்வெண்களாகவும் குறிக்கப்படும். பாகுபாடு செய்யும்போது ஒரேபடித்தான நிகழ்வெண்களைச் சேர்த்தமைத்தல் சிறப்பானதாகும். Nஐ (உறுப்புகளின் மொத்த எண்ணிக்கை) 0-படி நிகழ்வெண் எனக் கொண்டு, இரண்டு, மூன்று பண்புகளின் அடிப்படையில் பாகுபாடு செய்யப்பட்ட முழுமைத் தொகுதியை அடியிற்காணும் முறைப்படி வரிசைப்படுத்தலாம்.

இரண்டு பண்புகள் அமையப்பெற்ற முழுமைத் தொகுதி

0-படி	......	N-முழுமைத் தொகுதியின் மொத்த எண்ணிக்கை
1-முதற்படி	......	(A) (α) (B) (β)
2-இரண்டாம்படி	......	(AB) $(A\beta)$ (αB) $(\alpha\beta)$

மூன்று பண்புகள் அமையப் பெற்ற முழுமைத் தொகுதி

0-படி	......	N-முழுமைத் தொகுதியின் மொத்த எண்ணிக்கை
1-முதற்படி	......	(A) (α) (B) (β) (C) (γ)
2-இரண்டாம்படி	......	(AB) (AC) (BC) $(A\beta)$ $(A\gamma)$ $(B\gamma)$ (αB) (αC) (βC) $(\alpha\beta)$ $(\alpha\gamma)$ $(\beta\gamma)$
3-மூன்றும்படி	......	(ABC) $(AB\gamma)$ (αBC) $(A\beta C)$ $(A\beta\gamma)$ $(\alpha B\gamma)$ $(\alpha\beta\gamma)$ $(\alpha\beta\gamma)$

ண்பு நிகழ்வெண்களின் அமைப்பு முறை

முழுமைத் தொகுதியின் மொத்தம் $= N$

A-பண்பு பெற்ற உறுப்புகள் $= (A)$

A-பண்பு பெறாத உறுப்புகள் $= (\alpha)$

$$(A) + (\alpha) = N$$

A, α ஒன்றுக்கொன்று எதிரான பண்புகளாதலின், முழுமைத் தொகுதியில் உள்ள உறுப்புகள் A பண்பு அல்லது α பண்பு (A — அமையப்பெறாதது) ஆகிய இரண்டில் ஏதாவதொன்றைப் பெற்றாக வேண்டும்.

இதே போல்

$$(B) + (\beta) = N$$

$$(C) + (\gamma) = N$$

எனவும் அமைவதைக் காணலாம்.

இந் நியதி 'பொருத்தமான இரு முதற்படி நிகழ்வெண்களின் கூட்டுத்தொகை மொத்த உறுப்புகளுக்குச் சமமாகும்' என்பதைத் தெளிவுபடுத்துகிறது.

(AB) என்பது A பண்பு உடையனவற்றில் B பண்பு பெற்றவற்றின் எண்ணிக்கையையும்,

$(A\beta)$ என்பது A பண்பு உடையவற்றில் β பண்பு பெற்றவற்றின் (B பண்பு பெறாதவை) எண்ணிக்கையையும் குறிக்கும்.

$$\therefore\ (AB) + (A\beta) = (A)$$

இதேபோல்

$$(\alpha B) + (\alpha\beta) = (\alpha)$$

$$(AB) + (A\beta) + (\alpha B) + (\alpha\beta) = (A) + (\alpha) = N$$

மூன்று பண்புகள் அமையுமிடத்து (AB) என்பது C பண்பு பெற்றவை என்றும், C பண்பு பெறாதவை என்றும் (γ) பிரிக்கப்படும்.

$$(AB) = (ABC) = (AB\gamma)$$

இதேபோல்

$$(A\beta) = (A\beta C) + (A\beta\gamma)$$

$$(\alpha B) = (\alpha BC) + (\alpha B\gamma)$$

$$(\alpha\beta) = (\alpha\beta C) + (\alpha\beta\gamma)$$

$$\therefore\ (ABC) + (AB\gamma) + (A\beta C) + (A\beta\gamma) + (\alpha BC) + (\alpha B\gamma) + (\alpha\beta C) + (\alpha\beta\gamma)$$

$$= (AB) + (A\beta) + (\alpha B) + (\alpha\beta)$$

$$= (A) + (\alpha) = N$$

எனவே, 'கொடுக்கப்பட்ட முழுமைத் தொகுதிக்கான உயர் படி (higher order) நிகழ்வெண்களிலிருந்து, நேரிடை மற்றும் எதிரிடைப் பண்புகளுக்கான நிகழ்வெண்களைப் பெறலாம்' என்பது தெரியவருகிறது.

காட்டு :

இரு பண்புகளின் அடிப்படையில் பிரிக்கப்பட்ட முழுமைத் தொகுதிக்கான உயர்படி நிகழ்வெண்கள்,

$$(AB) = 50 \qquad (\alpha B) = 40$$

$$(A\beta) = 25 \qquad (\alpha\beta) = 20$$

எனக் கொடுக்கப்பட்டுள்ளன. முதற்படி நிகழ்வெண்களையும், மொத்த உறுப்புகளையும் கணக்கிடுவோம்.

(A), (B), (α), (β) என்பன நாம் கணக்கிடவேண்டிய முதற்படி நிகழ்வெண்கள்.

முதற்படி நிகழ்வெண்கள், அதே எழுத்தினைப் (குறியீட்டை) பெற்ற இரு இரண்டாம்படி நிகழ்வெண்களின் கூட்டுத்தொகைக்குச் சமமாகும். இதன் துணைகொண்டு முதற்படி நிகழ்வெண்கள் அனைத்தும் கணக்கிடப்படுகின்றன.

$$(A) = (AB) + (A\beta) = 50 + 25 = 75$$

$$(B) = (AB) + (\alpha B) = 50 + 40 = 90$$

$$(\alpha) = (\alpha B) + (\alpha\beta) = 40 + 20 = 60$$

$$(\beta) = (A\beta) + (\alpha\beta) = 25 + 20 = 45$$

மொத்த உறுப்புகள் :

$$(A) + (\alpha) = 75 + 60 = 135 = N$$

$$(B) + (\beta) = 90 + 45 = 135 = N$$

இரண்டாம்படி நிகழ்வெண்களிலிருந்து முதற்படி நிகழ்வெண்களைக் கணக்கிட அடியிற்காணும் பட்டியல் மிகவும் உதவியாக உள்ளது.

	B	β	
(A)	(AB) 50	$(A\beta)$ 25	(A) 75
α	(αB) 40	$(\alpha\beta)$ 20	(α) 60
	(B) 90	(β) 45	N 135

அறிந்த நிகழ்வெண்களை 'அதற்குச் சரியான அறைகளில் அமைத்து நமக்குத் தேவையான நிகழ்வெண்களைப் பெறுகிறோம்.

காட்டு :

மூன்று பண்புகளின் அடிப்படையில் பிரிக்கப்பட்ட முழுமைத் தொகுதிக்கான உயர்படி நிகழ்வெண்கள் கொடுக்கப்பட்டுள்ளன. இதற்கான இரண்டாம்படி மற்றும் முதற்படி நிகழ்வெண்களைக் கணக்கிடுவோம்.

$(ABC) = 75 \qquad (\alpha BC) = 98$

$(AB\gamma) = 310 \qquad (\alpha B\gamma) = 702$

$(A\beta C) = 106 \qquad (\alpha\beta C) = 74$

$(A\beta\gamma) = 489 \qquad (\alpha\beta\gamma) = 8415$

இரண்டாம்படி நிகழ்வெண்கள் :

$$(AB) = (ABC) + (AB\gamma) = 75 + 310 = 385$$
$$(AC) = (ABC) + (A\beta C) = 75 + 106 = 181$$
$$(BC) = (ABC) + (\alpha BC) = 75 + 98 = 173$$
$$(A\beta) = (A\beta C) + (A\beta\gamma) = 106 + 489 = 595$$

$$(A\gamma) = (AB\gamma) + (A\beta\gamma) = 310 + 489 = 799$$
$$(B\gamma) = (AB\gamma) + (\alpha\beta\gamma) = 310 + 702 = 1012$$
$$(\alpha B) = (\alpha BC) + (\alpha B\gamma) = 98 + 702 = 800$$
$$(\alpha C) = (\alpha BC) + (\alpha\beta C) = 98 + 74 = 172$$
$$(\beta C) = (A\beta C) + (\alpha\beta C) = 106 + 74 = 180$$
$$(\alpha\beta) = (\alpha\beta C) + (\alpha\beta\gamma) = 74 + 8415 = 8489$$
$$(\alpha\gamma) = (\alpha B\gamma) + (\alpha\beta\gamma) = 702 + 8415 = 9117$$
$$(\beta\gamma) = (A\beta\gamma) + (\alpha\beta\gamma) = 310 + 8415 = 8725$$

முதற்படி நிகழ்வெண்கள் :

$$(A) = (AB) + (A\beta) = 385 + 595 = 980$$
$$(B) = (AB) + (\alpha B) = 385 + 800 = 1185$$
$$(C) = (AC) + (\alpha C) = 181 + 172 = 353$$
$$(\alpha) = (\alpha B) + (\alpha\beta) = 800 + 8489 = 9289$$
$$(\beta) = (A\beta) + (\alpha\beta) = 595 + 8489 = 9084$$
$$(\gamma) = (A\gamma) + (\alpha\gamma) = 799 + 9117 = 9916$$

மொத்த உறுப்புகள் :

$$(A) + (\alpha) = 980 + 9289 = 10269$$
$$(B) + (\beta) = 1185 + 9084 = 10269$$
$$(C) + (\gamma) = 353 + 9916 = 10269$$

கொடுக்கப்பட்ட உயர்படி நிகழ்வெண்களிலிருந்து மற்றப் பண்புகளுக்கான நிகழ்வெண்களைக் கணக்கிடும் முறையைக் கண்டோம். முழுமைத் தொகுதியின் நேரிடைப் பண்புகளுக்கான நிகழ்வெண்களிலிருந்து முழுமைத் தொகுதி பிரிக்கப்படும் மற்றப் பண்புகளுக்கான நிகழ்வெண்களைக் கணக்கிடும் முறையைப் பார்ப்போம்.

முழுமைத் தொகுதியில் A-பண்பு பெற்றவர்களைக் கணக்கிட A ஐ ஓர் இயக்கும் N கருவியாகப் (operator) பயன்படுத்தி $A \cdot N = (A)$ எனக் குறிப்பிடுவது பழக்கம்.

$$(A) = A \cdot N$$

இதேபோல்,

$$(\alpha) = \alpha \cdot N$$
$$(A) + (\alpha) = A \cdot N + \alpha \cdot N$$
$$= (A + \alpha) \cdot N$$

$$N = (A + \alpha) \cdot N$$
$$A + \alpha = 1$$
$$\alpha = 1 - A$$

இக் குறியீட்டுச் சமன்பாட்டைப் (Symbolic equivalance of attribute) பயன்படுத்தி எதிரிடைப் பண்புகளுக்கான நிகழ்வெண் களை நேரிடைப் பண்புகளின் நிகழ்வெண்களால் கணக்கிட முடிவதை உணரலாம்.

$$\begin{aligned}
(A\beta) &= A\beta \cdot N \\
&= A \cdot (1 - B) \cdot N \\
&= \overline{A - AB} \cdot N \\
&= A \cdot N - AB \cdot N \\
&= (A) - (AB)
\end{aligned}$$

$$\begin{aligned}
(\alpha\beta) &= \alpha\beta \cdot N \\
&= (1 - A)(1 - B) \cdot N \\
&= \overline{1 - A - B + AB} \cdot N \\
&= N - A \cdot N - B \cdot N + AB \cdot N \\
&= N - (A) - (AB) + (AB)
\end{aligned}$$

இதேபோல்

$$\begin{aligned}
(\alpha\beta C) &= \alpha\beta C \cdot N \\
&= (1-A)(1-B) \cdot C \cdot N \\
&= \overline{1-A-B+AB} \cdot C \cdot N \\
&= \overline{C - AC - BC + ABC} \cdot N \\
&= (C) - (AC) - (BC) + (ABC)
\end{aligned}$$

$$\begin{aligned}
(\alpha\beta\gamma) &= \alpha\beta\gamma \cdot N \\
&= (1-A)(1-B)(1-C) \cdot N \\
&= (1-A-B+AB)(1-C) \cdot N. \\
&= \overline{1-A-B+AB-C+AC+BC-ABC} \cdot N \\
&= N - A \cdot N - B \cdot N - AB \cdot N - C \cdot N + AC \cdot N \\
&\quad + BC \cdot N - ABC \cdot N \\
&= N - (A) - (B) - (C) + (AB) + (BC) + (CA) \\
&\quad - (ABC)
\end{aligned}$$

காட்டு :

கொடுக்கப்பட்ட நேரிடை நிகழ்வெண்களின் உதவியால் முழுமைத் தொகுதிக்கான மற்றப் பண்புகளின் நிகழ்வெண்களைக் கணக்கிடு.

$$(A) = 80 \qquad (B) = 100$$

$$(AB) = 70 \qquad N = 250$$

இரு பண்புகளின் அடிப்படையில் பிரிக்கப்பட்டுள்ள இம் முழுமைத் தொகுதியில் கணக்கிடப்படவேண்டிய நிகழ்வெண்களுக்கான மற்றப் பண்புகள் $\alpha, \beta, (\alpha B), A\beta, \alpha\beta$ என்பனவாகும்.

$$\begin{aligned}
(\alpha) &= \alpha \cdot N \\
&= (1-A) \cdot N \\
&= N - A \cdot N \\
&= N - (A) \\
&= 250 - 80 = 170
\end{aligned}$$

$$\begin{aligned}
(\beta) &= \beta \cdot N \\
&= (1-B) \cdot N \\
&= N - B \cdot N \\
&= N - (B) \\
&= 250 - 100 = 150
\end{aligned}$$

$$\begin{aligned}
(\alpha B) &= \alpha B \cdot N \\
&= (1-A) \cdot B \cdot N \\
&= \overline{B - AB} \cdot N \\
&= (B) - (AB) \\
&= 100 - 70 = 30
\end{aligned}$$

$$\begin{aligned}
(A\beta) &= A\beta \cdot N \\
&= A\,(1-B)\; N \\
&= \overline{A - AB} \cdot N \\
&= (A) - (AB) \\
&= 80 - 70 = 10
\end{aligned}$$

$$
\begin{aligned}
(\alpha\beta) &= \alpha\beta \cdot N \\
&= (1-A)\ (1-B) \cdot N \\
&= \overline{1-A-B+AB} \cdot N \\
&= N-A \cdot N-B \cdot N + AB \cdot N \\
&= N-(A)-(B)+(AB) \\
&= 250-80-100+70 = 140
\end{aligned}
$$

காட்டு :

மூன்று பண்புகளின் அடிப்படையில் பிரிக்கப்பட்ட நேரிணைப் பண்புகளின் நிகழ்வெண் உதவியால் முழுமைத் தொகுதிக்கான மற்றப் பண்புகளின் நிகழ்வெண்களைக் கணக்கிடுவோம்.

$$
\begin{aligned}
N &= 12000 & (AB) &= 453 \\
(A) &= 977 & (AC) &= 284 \\
(B) &= 1185 & (BC) &= 250 \\
(C) &= 596 & (ABC) &= 127
\end{aligned}
$$

நம்மால் கணக்கிடப்படவேண்டிய நிகழ்வெண்களுக்கான பண்புகள்

$$
\begin{array}{cccccc}
\alpha, & \beta, & \gamma, & \alpha B, & \alpha\beta, & A\gamma \\
\alpha C, & \alpha\gamma, & B\gamma, & \beta C, & \beta\gamma, & AB\gamma \\
A\beta C, & \alpha BC, & A\beta\gamma, & \alpha B\gamma, & \alpha\beta C, & \alpha\beta\gamma
\end{array}
$$

$$
\begin{aligned}
(\alpha) &= N-(A) = 12000 - 977 = 11023 \\
(\beta) &= N-(B) = 12000 - 1185 = 10815 \\
(C) &= N-(C) = 12000 - 596 = 11404
\end{aligned}
$$

$$
\begin{aligned}
(AB\gamma) &= (AB) - (ABC) = 453 - 127 = 326 \\
(A\beta C) &= (AC) - (ABC) = 284 - 127 = 157 \\
(\alpha BC) &= (BC) - (ABC) = 250 - 127 = 123 \\
(A\beta\gamma) &= (A\beta) - (A\beta C) = (A) - (AB) - (A\beta C) \\
&= 977 - 453 - 157 = 367 \\
(\alpha B\gamma) &= (\alpha B) - (\alpha BC) = (B)\,(AB) - (\alpha BC) \\
&= 1185-453-123 = 609 \\
(\alpha\beta C) &= (\beta C)-(A\beta C) = (C)-(BC)-(A\beta C) \\
&= 596-250-157 = 189
\end{aligned}
$$

$$(A\beta) = (A\beta C) + (A\beta\gamma) = 157 + 367 = 521$$
$$(\alpha B) = (\alpha BC) + (\alpha B\gamma) = 123 + 609 = 732$$
$$(A\gamma) = (AB\gamma) + (A\beta\gamma) = 326 + 367 = 693$$
$$(\alpha C) = (\alpha BC) + (\alpha\beta C) = 123 + 189 = 312$$
$$(B\gamma) = (AB\gamma) + (\alpha B\gamma) = 326 + 609 = 935$$
$$(\beta C) = (ABC) + (\alpha\beta C) = 157 + 189 = 935$$
$$(\alpha\beta\gamma) = (\alpha\beta) - (\alpha\beta C) = (\beta) - (A\beta) - (\alpha\beta C)$$
$$= 10815 - 524 - 189 = 10102$$
$$(\alpha\beta) = (\alpha\beta C) + (\alpha\beta\gamma) = 189 + 10102 = 10291$$
$$(\alpha\gamma) = (\alpha B\gamma) + (\alpha\beta\gamma) = 609 + 10102 = 10711$$
$$(\beta\gamma) = (A\beta\gamma) + (\alpha\beta\gamma) = 367 + 10102 = 10469$$

கணக்கியல் எதிர்பார்ப்பு (Mathematical expectation)

ஒரு நிகழ்ச்சி அல்லது பண்பிற்கான கணக்கியல் எதிர்பார்ப்பு அப் பண்பு அல்லது நிகழ்ச்சி நிகழ்வதற்கான நிகழ்திறனை (probability) மொத்த உறுப்புகளால் பெருக்கிவரும் மதிப்பிற்குச் சமமாகும். பத்து நாணயங்களைச் சுண்டும்போது கணக்கியலாக எதிர்பார்க்கும் தலைகள் $\frac{1}{2} \times 10 = 5$ என அமையும். ஒரு நாணயத்தைச் சுண்டும்போது கிடைக்கும் தலைவிழுவதற்கான நிகழ்திறனைச் சுண்டிய மொத்த நாணயங்களால் பெருக்கிக் கணக்கியல் எதிர்பார்ப்புக் குறிக்கப்பட்டது. பத்து நாணயங்களைச் சுண்டும் போது நமக்குக் கிடைக்கவேண்டிய தலைகளின் எண்ணிக்கையையே இது குறிக்கிறது. சோதனை முடிவுகள் இதனை ஒத்தும் மாறுபட்டும் இருக்கலாம். சோதனை முடிவுகளைக் கணக்கியல் எதிர்பார்ப்புடன் ஒப்பிட்டுப் புள்ளியியலில் பெரும்பாலான முடிவுகள் செய்யப்படுகின்றன.

இயைபற்ற தன்மைக்கான நியதி (Cirterion for independence of attributes)

சோதனை முடிவுகளும், கணக்கியலாக எதிர்பார்க்கும் முடிவுகளும் சமமாக இருப்பின், இரு பண்புகளும் இயைபற்றன என்று முடிவைச் சொல்லுகிறோம்.

காட்டு :

$$(A) = 60 \qquad (B) = 90$$
$$(AB) = 54 \qquad N = 100$$

A, B ஆகிய பண்புகளின் இயைபைச் சோதிப்போம்.

(AB)-யின் சோதனை முடிவு $= 54$

A, B ஆகிய இரு பண்புகளும் சேர்ந்து அமையப்பெற்ற உறுப்புகள் 54 என்பதை இது குறிக்கிறது.

$$A\text{-பண்பு நிகழ்வதற்கான நிகழ்திறன்} = \frac{(A)}{N}$$

$$B\text{-பண்பு நிகழ்வதற்கான நிகழ்திறன்} = \frac{(B)}{N}.$$

$$\left.\begin{array}{l}A, B \text{ ஆகிய இரு பண்புகளும் சேர்ந்து} \\ \text{நிகழ்வதற்கான நிகழ்திறன்}\end{array}\right\} = \frac{(A)}{N} \cdot \frac{(B)}{N}$$

$$\left.\begin{array}{l}A, B \text{ இரு பண்புகளும் சேர்ந்து நிகழ்வதற்கான} \\ \text{கணக்கியல் எதிர்பார்ப்பு}\end{array}\right\} = P(A) \cdot P(B) \cdot N$$

$$\Big[(A, B \text{ இரு பண்புகளும் சேர்ந்து நிகழ்வதற்கான நிகழ்திறன்}) \times (\text{மொத்த உறுப்புகள்})\Big]$$

$$= \frac{(A)}{} \times \frac{(B)}{N} \times N = \frac{(A)(B)}{N}$$

$$= \frac{60 \times 90}{100} = 54$$

சோதனை முடிவுகளும், கணக்கியல் எதிர்பார்ப்பும் சமமாக இருப்பதால் A, B ஆகிய இரு பண்புகளும் இயைபற்றன என்ற முடிவைச் சொல்கிறோம்.

இயைபு (association) என்ற சொல்லுக்கு நடைமுறையில் கொடுக்கப்படும் பொருளினின்றும் சற்று மாறுபட்ட பொருளில் புள்ளியியல் பிரிவில் உபயோகப்படுத்தப்படுவதை உணரவேண்டியது அவசியமாகும். முழுமைத் தொகுதியில் இரு பண்புகள் சேர்ந்தமைந்த உறுப்புகள் கிடைக்கப்பெற்ற நிலையில் அவை இயைபுடையன என்ற முடிவைச் சொல்வது நடைமுறைப் பழக்கமாகும். இயைபற்ற நிலையில் அவ் விருபண்புகளும் அமையப்பெறும் உறுப்புகளின் எண்ணிக்கையைக் காட்டிலும் அதிகமான உறுப்புகளிடத்து இரு பண்புகளும் அமையப்பெற்ற விடத்தே இரு பண்புகளும் இயைபுடையன என்ற முடிவைக் கூறுவது புள்ளியியலில் இயைபு (association) என்ற சொல்லுக்குக் கொடுக்கப்படும் பொருளாகும்.

சோதனை முடிவுகள், கணக்கியலாக எதிர்பார்க்கும் முடிவுகளைக் காட்டிலும் அதிகமாக இருப்பின், இரு பண்புகளுக்கும் நேரிடை இயைபு (possitive association) உண்டு எனச் சொல்கிறோம்.

$(AB) > \frac{(A)(B)}{N}$ என்ற நிலையில் A என்ற பண்பும், B என்ற பண்பும் நேரிடை இயைபுடையன என்ற முடிவைச் சொல்கிறோம்.

காட்டு :

α, β ஆகிய இரு பண்புகளுக்கான இயைபின் தன்மையைப் பார்ப்போம்.

கொடுக்கப்பட்டது :

$$(\alpha) = 150 \qquad (\alpha\beta) = 70$$
$$(\beta) = 210 \qquad N = 500$$

α, β ஆகிய இரு பண்புகளும் சேர்ந்து நிகழ்வதற்கான சோதனை முடிவு $= (\alpha\beta) = 70$

α, β ஆகிய இரு பண்புகளும் சேர்ந்து நிகழ்வதற்கான கணக்கியல் எதிர்பார்ப்பு $= \frac{(\alpha)\cdot(\beta)}{N} = \frac{150 \times 210}{500} = 63$

சோதனை முடிவு $[(\alpha\beta) = 70]$ கணக்கியலாக எதிர்பார்க்கும் முடிவைவிட அதிகமாக இருப்பதால் α, β என்ற இரு பண்புகளும் நேரிடை இயைபுடையன (passitively associated) எனச் சொல்கிறோம்.

சோதனை முடிவுகள் கணக்கியலாக எதிர்பார்க்கும் முடிவுகளைக் காட்டிலும் குறைவாக இருப்பின், இரு பண்புகளுக்குமான இயைபு எதிரிடை இயைபு (negative assosiation) எனக் கொள்கிறோம். அதாவது $(AB) < \frac{(A)(B)}{N}$ என்ற நிலையில் A, B என்ற இரு பண்புகளும் எதிரிடை இயைபுடையன எனச் சொல்கிறோம். எதிரிடை இயைபை இயைபற்ற தன்மை (dissociation) எனவும் சொல்லலாம்.

காட்டு :

A, β ஆகிய பண்புகளின் இயைபின் தன்மையைக் கணக்கிடுவோம்.

$$(A) = 240 \qquad (\beta) = 300$$

சோதனை முடிவுகள் :

$$(A\beta) = 80 \qquad N = 800$$

A, β ஆகிய இரு பண்புகளும் சேர்ந்து நிகழ்ந்த சோதனை முடிவு } = 80

A, β ஆகிய இரு பண்புகளும் சேர்ந்து நிகழ்வதற்கான கணக்கியல் எதிர்பார்ப்பு } $= \frac{(A)(\beta)}{N} = \frac{240 \times 300}{800} = 90$

சோதனை முடிவு, கணக்கியலாக எதிர்பார்க்கும் முடிவைக் காட்டிலும் குறைவாக இருப்பதால் A, β ஆகிய இரு பண்புகளும் எதிரிடை இயைபுடையன என்றோ அல்லது இயைபற்றன என்றோ சொல்கிறோம்.

இரு பண்புகளுக்கான இயைபின் தன்மையை $(AB) = \frac{(A)(B)}{N}$ என்ற கட்டளை விதிப்படி குறிப்பிடுவது கணக்கியல் முறைப்படி சரியாக அமைந்தாலும், நடைமுறையில் வேறு சில யுக்திகளும் உபயோகப்படுத்தப்படுகின்றன. A உடன் அமைந்த B, A உடன் அமைந்த βC (B-அல்லாதது) ஆகியவற்றின் விகிதசமம் அல்லது சதவிகிதங்களைக் கணக்கிட்டு, A, B க்கான இயைபின் தன்மையைக் கூறுவதும் உண்டு.

காட்டு :

அடியிற்காணும் சோதனை முடிவிலிருந்து தடுப்பூசி போடுவதற்கும் காலரா (cholera) வராமலிருப்பதற்கும் உள்ள இயைபின் தன்மையைப் பார்ப்போம்.

	வியாதியால் பாதிக்கப் படாதவர்கள்	வியாதியால் பாதிக்கப் பட்டவர்கள்	மொத்தம்
தடுப்பூசி போட்டவர்கள்	276	3	279
தடுப்பூசி போடாதவர்கள்	473	66	539
ாத்தம்	749	69	818

தடுப்பூசி போட்டவர்களில் வியாதியால் பாதிக்கப்படாத வர்களின் சதவீதம்	=	$\frac{276}{279} \times 100 = 98{\cdot}9$
தடுப்பூசி போடாதவர்களில் வியாதியால் பாதிக்கப்படாத வர்களின் சதவீதம்	=	$\frac{473}{539} \times 100 = 87{\cdot}8$
தடுப்பூசி போட்டவர்களில் வியாதியால் பாதிக்கப்பட்ட வர்கள்	=	$\frac{3}{279} \times 100 = 1{\cdot}1\%$
தடுப்பூசி போடாதவர்களில் வியாதியால் பாதிக்கப்படாத வர்கள்	=	$\frac{66}{539} \times 100 = 12{\cdot}2\%$

இரண்டு முறையான ஒப்பீடும், தடுப்பூசி போடுவதும் வியாதியால் பாதிக்கப்படாமையும் நேரிடை இயைபுடையன (positively associatied) என்பதை உறுதிப்படுத்துகிறது.

இயைபு கெழு (Co-efficient of Association)

இயைபற்ற தன்மைக்கான நியதியின் உதவியால் இரு பண்புகள் இயைபற்றனவா, எதிரிடை அல்லது நேரிடை இயைபுடையனவா என்பதற்கான முடிவைச் சொல்ல முடியும் என்பதை அறிந்தோம். இக் கட்டளைப்படி இரு பண்புகள் எத்தகைய இயைபைப் பெற்றுள்ளன என்பதைச் சொல்ல முடியுமேயன்றி குறிப்பிட்ட பண்புகள் எந்த அளவினதான இயைபைப் பெற்றுள்ளன என்பதைக் கூற முடியாத நிலையில் உள்ளோம். இயையின் தன்மையை அளவிட்டுக் கூறுவதற்கும், இரு மாதிரிகளில் அமைந்த பண்புகளுக்கான இயைபின் தன்மையை ஒப்பிட்டுப் பார்க்கவும் இயைபு கெழுக்களைக் கணக்கிட வேண்டிய அவசியம் ஏற்படுகிறது. யூல் (Yule) என்பவரின் இயைபுகெழு இரு பண்புகளுக்கான இயைபின் அளவைக் கணக்கிட உதவுகிறது.

$$\text{இயைபு கெழு} = Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha\beta)}{(AB)(\alpha\beta) + (A\beta)(\alpha\beta)}.$$

யூலின் இயைபு கெழு எளிதாகக் கணக்கிட முடிவதாயும் புரிந்துகொள்வதற்கு இலகுவாயும் அமைந்துள்ளது. இரு பண்புகளும் இயைபற்ற நிலையில் உள்ளவிடத்து யூலின் இயைபு கெழு பூச்சிய மதிப்பைக் கொடுக்கும். இரு பண்புகளும் முழுக்க நேரிடை இயைபைப் பெற்ற நிலையில் $+1$ என்ற மதிப்பையும், முழுக்க எதிரிடை இயைபைப் பெற்ற நிலையில் -1 என்ற மதிப்பையும் கொடுக்கும்.

$Q = 0$ என வருவதற்கு

$(AB)(\alpha\beta) - (\alpha B)(A\beta) = 0$ என்றமைய வேண்டும்.

$Q = -1$ என வருவதற்கு

$(AB)(\alpha\beta) = 0$ என்றமைய வேண்டும்.

அதாவது $(AB) = 0$ அல்லது $(\alpha B) = 0$ என்றிருக்க வேண்டும்.

$Q = +1$ என வருவதற்கு

$(A\beta)(\alpha B) = 0$ என்றமைய வேண்டும்.

அதாவது $(A\beta) = 0$ அல்லது $(\alpha B) = 0$ என்றமைய வேண்டும்.

$(A) = (AB) + (A\beta)$ என்பது

$(A) = (AB)$ என்றாகிறது.

'A'-பண்புடையவர்கள் அனைவரும் 'B'-பண்பையும் பெற்றவர்கள் என்பதை இது காட்டுகிறது.

− 1-லிருந்து + 1 வரை அமையப்பெறும் இவ் இயைபு கெழுவின் மதிப்பிற்குத் தக்கவாறு இரு பண்புகளுக்கான இயைபின் தன்மை விளக்கப்படும். இரு பண்புகளுக்கான இயைபு மிகக் குறைந்த நிலையில் அதாவது (AB) என்பது $\frac{(A) \times (B)}{N}$ -லிருந்து குறிப்பிடத்தக்க வித்தியாசத்தைப் பெறுமலிருக்கும் போது, இரு பண்புகளுக்கான இயைபு ஒரு நிலையான மதிப்பைப் பெற்றதாகச் சொல்ல முடியாது. நம்மால் குறிப்பிட முடியாத காரணங்களினாலோ அல்லது மாதிரி எடுத்தலின் ஏற்ற இறக்கங்களினாலோ சிறு வித்தியாசம் ஏற்படலாம் என்பதையே இது காட்டுகிறது. திரும்பத் திரும்ப எடுக்கப்படும் மாதிரிகளின் உதவியினாலேயே இத்தகைய நிலையில் ஒரு முடிவான கருத்தைச் சொல்ல முடியும்.

காட்டு :

	குணமானவர்கள்	குணமாகாதவர்கள்	மொத்தம்
நோய் நீக்க மருந்தை உபயோகித்தவர்கள்	35	50	85
நோய் நீக்க மருந்தை உபயோகிக்காதவர்கள்	55	260	315
மொத்தம்	90	310	400

பண்புக் குறியீடுகளில் குறிப்பிடும்போது,

நோய் நீக்க மருந்தை உபயோகித்தவர்கள் = A

நோய் நீக்க மருந்தை உபயோகிக்காதவர்கள் = α

வியாதியிலிருந்து குணமானவர்கள் = B

வியாதியிலிருந்து குணமாகாதவர்கள் = β

என அமையும்,

இதனையே 2×2 இணைப்பட்டியலாக (2×2 contingency table) அமைக்கும் போது

	B	β	
A	(AB) 35	$(A\beta)$ 50	(A) 85
α	(αB) 55	$(\alpha\beta)$ 260	(α) 315
	(B) 90	(β) 310	(N) 400

என அமையும்.

$$Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) + (A\beta)(\alpha B)}$$

$$= \frac{(35 \times 260) - (55 \times 50)}{(35 \times 260) + (55 \times 50)} = + \frac{6350}{11850}$$

$$= + 0{\cdot}58$$

மருந்தைப் பயன்படுத்துவதும், நோயினின்று விடுபடுவதும் நேரிடை இயைபுடையன (positively associated) என்ற முடிவை இது உறுதிப்படுத்துகிறது.

காட்டு :

தந்தைமார்களின் கண் நிறத்திற்கும், அவர்தம் பிள்ளைகளின் கண் நிறத்திற்கும் உள்ள இயைபின் தன்மையை அடுத்த பக்கத்திற் காணும் முடிவுகளிலிருந்து கணக்கிடுவோம்.

		பிள்ளைகளின் கண் நிறம்		
		வெளிச்ச மற்றன B	வெளிச்ச முடையன β	
தந்தைமார்களின் கண் நிறம்	வெளிச்சமற்றன A	(AB) 230	$(A\beta)$ 150	(A) 380
	வெளிச்ச முடையன α	(αB) 150	$(\alpha\beta)$ 470	(α) 620
		(B) 380	(β) 620	N 1000

$$Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) + (A\beta)(\alpha B)}$$

$$= \frac{(230)(470) - (150)(150)}{(230)(470) + (150)(150)}$$

$$= +\,0{\cdot}655$$

தந்தைமார்களின் கண் நிறமும், அவர்தம் பிள்ளைகளின் கண் நிறமும் நேரிடை இயைபுடையன என்ற முடிவைச் சொல்கிறோம்.

பகுதி இயைபு (Partial Association)

இரு பண்புகளுக்கான இயைபினைக் கணக்கிடும் முறையைக் கண்டோம். மூன்று பண்புகளில் ஏதேனும் ஒரு பண்பு உடையவர் களிடையே மற்ற இரு பண்புகளுக்கு உள்ள நேரிடை அல்லது எதிரிடை இயைபு, பகுதி இயைபு (Partial association) எனப்படும். தடுப்பூசி போடுவதற்கும் அம்மையால் பாதிக்கப்படாமலிருப்ப தற்கும் நேரிடை இயைபுடையதாகக் கணக்கிட்டோம். படித்த வர்களாலும், நல்ல நிலையில் (well-to-do) உள்ளவர்களாலுமே தடுப்பூசி சந்தேகத்திற்கிடமின்றி ஏற்கப்படுவதை உணரும்போது, தடுப்பூசி போடுவதற்கும் அம்மை வராமலிருப்பதற்கும் உள்ள இயைபு நம் நாட்டைப் பொறுத்தவரை வேறொரு பண்பால்

இயைபுடைய நிலையை அடைவதை உணர முடிகிறது. *A, B* ஆகியவற்றின் இயைபு முழுமைத் தொகுதிக்கான கருத்தாக அமையாமல் *C* எனும் பண்பு அமையப்பெற்ற துணை முழுமைத் தொகுதிக்கு (sub-population) பொருந்துவதைப் பகுதி இயை புடையவை எனச் சொல்கிறோம். *A, B, C* என்ற மூன்று பண்பு களில் *C* என்ற பண்பை அடிப்படையாகக் கொண்டு, *A, B*-க் குள்ள இயைபற்ற தன்மையை

$$(ABC) = \frac{(AC)(BC)}{(C)}$$

என்ற நியதியைக் கொண்டு குறிப்பிடுகிறோம்.

$$(ABC) > \frac{(AC)\,(BC)}{(C)} \quad , \quad (ABC) < \frac{(AC)\,(BC)}{(C)}$$

என்ற நியதிகள், *C* என்ற பண்பின் அடிப்படையில் *B, C*-க்குள்ள நேரிடை, எதிரிடை இயைபைக் குறிக்கும் நியதிகளாகும்.

பகுதி இயைபுடைய நிலை, 'ஒருதலை இயைபு', 'ஒருசிறை இயைபு' என்வும் சொல்லப்படும். பகுதி இயைபுடைய நிலை மிகத் தெளிவாக வரையறுக்கப்பட வேண்டும். துணை முழுமைத் தொகுதிக்குக் கூறப்படும் கருத்துகள், முழுமைத் தொகுதிக்கும் பொருந்தாதாகையால், இந் நிலை தெளிவுபடுத்தப்பட்டாலன்றி இயைபு கெழுவின் விளக்கம் நிலையற்ற தன்மையை அடைந்து விடும். '*C*' என்ற பண்பின் அடிப்படையில் இயைபுடைய இரு பண்புகள், '*D*' அல்லது '*E*' என்ற பண்புகளின் அடிப்படையிலும் இயைபுடை நிலையில் இருக்கவேண்டிய அவசியமில்லாமையையே இது உணர்த்துகிறது.

பகுதி இயைபிற்கான நியதிகள் ஒரு பண்பின் அடிப்படையில் மற்ற இரு பண்புகளுக்கான இயைபின் தன்மையைமட்டும் தெளிவுபடுத்தும். இயைபு கெழுவைப் போலவே பகுதி இயைபு கெழுவையும் (co-efficient of partial association) கணக் கிட்டு ஒரு பண்பின் அடிப்படையில் இரு பண்புகளின் இயைபிற் கான நெருக்கத்தைக் கணக்கிட முடியும்.

$$\text{பகுதி இயைபு கெழு} = Q_{AB \cdot C} = \frac{(ABC)(\alpha\beta C) - (A\beta C)(\alpha BC)}{(ABC)(\alpha\beta C) + (A\beta C)(\alpha BC)}$$

காட்டு :

கொடுக்கப்பட்ட முடிவிலிருந்து *A, B* ஆகிய பண்புகளுக் கான இயைபையும், பகுதி இயைபையும் கணக்கிடுவோம்.

$$(A) = 348 \qquad (AB) = 330$$
$$(B) = 455 \qquad (BC) = 353 \qquad (ABC) = 300$$
$$(C) = 465 \qquad (AC) = 310 \qquad N = 600$$

இயைபு கெழு $= Q = \dfrac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) + (A\beta)(\alpha B)}$

கொடுக்கப்பட்ட நிகழ்வெண்களிலிருந்து $(\alpha\beta)$, $(A\beta)$, (αB) ஆகியவற்றிற்கான நிகழ்வெண்களைக் கணக்கிட்டு இயைபு கெழு கணக்கிடவேண்டி உள்ளது.

$$(\alpha\beta) = N-(A)-(B)+(AB)$$
$$= 600-348-455+330 = 127$$
$$(A\beta) = (A)-(AB) = 348-330 = 18$$
$$(\alpha B) = (B)-(AB) = 455-330 = 125$$
$$\alpha = \frac{(330)(127)-(18)(125)}{(330)(127)+(18)(125)} = 0{\cdot}950$$

பகுதி இயைபு கெழு $Q_{AB\cdot C} = \dfrac{(ABC)(\alpha\beta C)-(A\beta C)(\alpha BC)}{(ABC)(\alpha\beta C)+(A\beta C)(\alpha BC)}$

$$(\alpha\beta C) = (1-A)(1-B)\cdot C\cdot N$$
$$= (C)-(AC)-(BC)+(ABC)$$
$$= 465-310-353+300 = 102$$
$$(A\beta C) = A(1-B)\cdot C$$
$$= (AC)-(ABC)$$
$$= 310-300 = 10$$
$$(\alpha BC) = (1-A)\,B\cdot C\cdot N$$
$$= (BC)-(ABC)$$
$$= 353-300$$
$$= 53$$
$$Q_{AB\cdot C} = \frac{(300)(102)-(10)(53)}{(300)(102)+(10)(53)} = 0{\cdot}966$$

போலி இயைபு (Illusory Association)

உண்மையான இயைபிலா நிலையைப் போலி இயைபுடையவை எனச் சொல்கிறோம். C என்ற பண்பினால் ஏற்படும் A, B

பண்புகளுக்கான இயைபு சில இடங்களில் போலி இயைபாய் முடிகிறது. படிப்பும் [A], வேலையின் தன்மையும் [B], நல்ல வாழ்க்கைத் தரத்துடன் [C] தொடர்புடைய தன்மையைக் கொடுத்தாலும், முழுமைத் தொகுதியில் இது ஓர் இயைபற்ற நிலையைத்தான் கொடுக்கிறது. *A, B* ஆகிய பண்புகள் சரியாக வரையறுக்கப்படாதவிடத்தும், பண்புகளைப் பகுப்பதில் சோதனை யாளர்தம் தாராள மனப்பான்மையும் போலி இயைபை (Illusory association) உண்டாக்கலாம். இரு பண்புகளுக்கான இயைபின் நிலைமையைத் தெளிவுபடுத்தி ஒரு நிலையான உருவைக் கொடுக்க, போலி இயைபேற்படும் வழிகளைத் தவிர்த்துப் பண்புகள் பிரித்தமைக்கப்படுவது இன்றியமையாததாகும்.

இசைவுடைமை (Consistency)

ஒரு முழுமைத் தொகுதியின் அளவுகள் யாவும் ஒன்றுக்கொன்று முரண்படாமலிருப்பது அவசியமாகும். ஒன்றுக்கொன்று முரண்படாத நிலை இசைவுடைமை (consistent) எனச் சொல்லப்படுகிறது. பண்புகளின் இயைபிற்கான நெருக்கத்தைக் கணக்கிடுமுன், அவை முரண்படாதுள்ளனவா என்பதைச் சோதனை செய்வது அவசியமாகும். எந்தவொரு பண்பிற்குமான நிகழ்வெண் எதிரெண்ணுக (negative frequency) அமையாததை அறிவோம். யாருமே ஒரு பண்பைப் பெறுதவர்களாக இருக்கலாம். ஆனால், அந்தப் பண்பிற்கான நிகழ்வெண் எதிரெண்ணுக அமைய முடியாது. இப் பண்பை நியதியாகக்கொண்டு கொடுக்கப்பட்ட முடிவுகளுக்கான இசைவுத் தன்மை முடிவு செய்யப்படுகிறது. ‘எந்த உயர்படி நிகழ்வெண்ணும் (Ultimate frequency) எதிரெண்ணுக அமைய முடியாது’ என்ற நியதியே நமது சோதனையின் தன்மையாகும். கொடுக்கப்பட்ட அளவுகளுக்கான உயர்படி நிகழ்வெண்களைக் கணக்கிட்டு, எல்லா உயர்படி நிகழ்வெண்களும் நேரெண்ணுக அமைந்தால், அளவுகள் இசைவுடையன (consistent) என்றும், ஏதாவதோர் உயர்படி நிகழ்வெண் எதிரெண்ணுக அமைந்தாலும் அளவுகள் இசைவற்று உள்ளன (inconsistent) என்றும் முடிவு செய்கிறோம்.

காட்டு :

அடியிற்காணும் அளவுகளுக்கான இசைவுத் தன்மையைப் பார்ப்போம்.

$$N = 1000 \qquad (A\beta) = 350$$
$$(AB) = 200 \qquad (\alpha\beta) = 500$$

கொடுக்கப்பட்ட முடிவுகளில் தவறு வெளிப்படையாகத் தெரியாவிட்டாலும்,

$$(\alpha\beta) = N-(AB)-(A\beta)+(\alpha B)$$
$$= 1000 - 200 - 350 + 500 = -50$$

என அமைந்து, அளவுகளின் இசைவற்ற தன்மையைக் காட்டுகிறது.

காட்டு :

$(A\beta) = 414$ $(\alpha\beta) = 238$

$(\alpha B) = 260$ $(A) = 400$

$$(AB) = (A) - (A\beta) = 400-414 = -14$$

அளவுகளின் இசைவற்ற தன்மையைக் காட்டுகிறது.

காட்டு :

$N = 2000$ $(AB) = 676$

$(A) = 1754$ $(AC) = 286$

$(B) = 2172$ $(BC) = 220$

$(C) = 572$ $(ABC) = 114$

$$(\alpha\beta\gamma) = N - (A) - (B) - (C) + (AB) + (AC)+(BC) - (ABC)$$
$$= 2000-1754-2172-572+676+286+220-114$$
$$= -1430$$

இசைவற்றுள்ளன.

இணைப்பட்டியல் (Contingency table)

முழுமைத் தொகுதியை A, B என்ற இரு பண்புகளின் அடிப்படையில் பிரிக்கும்போது 2×2 இணைப்பட்டியலை அமைத்து, பண்புகளின் பகுப்பிற்கான நிகழ்வெண்களையும், இயைபுகெழுவினையும் கணக்கிடும் முறையைக் கண்டோம். A என்ற பண்பு $A_1, A_2 \ldots\ldots\ldots A_m$ எனவும், B என்ற பண்பு $B_2, B_1 \ldots\ldots\ldots B_n$ எனவும் பகுக்கப்படும்போது கிடைக்கும் $m \times n$ இணைப்பட்டியலுக்கான இயைபு கெழுவினைக் கணக்கிடும் முறையைப் பார்ப்போம். A என்ற பண்பினை $A_1, A_2 \ldots\ldots A_m$ என்ற m படுகைநிலை (row) வரிசைகளாகவும், B என்ற பண்பினை $B_2, B_1 \ldots\ldots B_n$ என்ற நிமிர்நிலை (colnmn) வரிசைகளாகவும் அமைக்கும்போது கிடைக்கும் பட்டியலே $m \times n$ இணைப்பட்டியலாகும் ($m \times n$ contingency table).

$m \times n$ இணைப்பட்டியல்

B \ A	B_1	B_2		B_s		B_n	
A_1	(A_1B_1)	(A_1B_2)		(A_1B_s)		A_1B_n	(A_1)
A_2	(A_2B_1)	(A_2B_2)		$(A_2B\ )$		A_2B_n	(A_2)
A_r	(A_rB_1)	(A_rB_2)		(A_rB_s)		A_rB_n	$(A\ )$
A_m	(A_mB_1)	(A_mB_2)		(A_mB_s)		A_mB_n	(A_m)
	(B_1)	(B_2)		(B_s)		(B_n)	N

$A_r B_s$ என்ற பண்பிற்கான நிகழ்வெண் A பண்பிற்கான 'r' படுகைநிலை வரிசை, B பண்பிற்கான 's' நிமிர்நிலை வரிசை ஆகியவற்றின் பொது அறையில் (common cell) குறிக்கப்படும். குறிப்பிட்ட பண்பிற்கான நிகழ்வெண் அப் பண்பு அமைந்த நிமிர் அல்லது படுகைநிலை வரிசைகளின் மொத்தத்திற்குச் சமமாகும். இவை யாவும் முதற்படி நிகழ்வெண்களாக (first order frequency) அமைவதை உணரலாம்.

இணைப்புக் கெழு (Co-efficient of Contingency)

இணைப்பட்டியலுக்கான இயைபின் அளவையே இணைப்புக் கெழு எனப்படும். இரு பண்புகளுக்கான சோதனை முடிவுகளும் கணக்கியலாக எதிர்பார்க்கும் முடிவுகளுக்குச் சமமாயின், அவை இயைபற்றன என்ற முடிவையும், அவற்றுக்குள்ள வித்தியாசத்தின் தன்மைக்கேற்ப இயைபின் தன்மை குறிப்பிடப்பட்டதையும் அறிவோம். இரு பண்புகளுக்கான கணக்கியல் எதிர்பார்ப்பு நிகழ்வெண், பண்பு நிகழ்வதற்கான நிகழ்திறனை மொத்த உறுப்பு

களின் எண்ணிக்கையால் பெருக்கிவரும் தொகைக்குச் சமம் என்பதை அறிவோம். இணைப்புப் பட்டியலில் உள்ள ஒவ்வோர் அறைக்கும் கணக்கியல் எதிர்பார்ப்பு நிகழ்வெண் கணக்கிடப் பட்டு இணைப்புக் கெழு கணிக்கப்படுகிறது.

ஓர் அறைக்கான நிகழ்திறத்தை மொத்த உறுப்புகளால் பெருக்கினல் அவ் வறைக்கான கணக்கியல் எதிர்பார்ப்பு நிகழ் வெண் கிடைக்கும். ($A\ B_S$) என்ற அறைக்கான கணக்கியல் எதிர்பார்ப்பு நிகழ்வெண் கணக்கிடும் முறையைப் பார்ப்போம். நிகழ்திறனனது சாதக நிகழ்ச்சிகளை மொத்த நிகழ்ச்சிகளால் வகுத்துவரும் தொகைக்குச் சமமாகும். $A_r\ B_S$ என்பது A_r என்ற பண்பும் B_S, என்ற பண்பும் சேர்ந்து நிகழ்வதைக் குறிக்கும். இதற்கான எதிர்பார்ப்பு நிகழ்வெண், இரு நிகழ்ச்சிகளும் (பண்பு களும்) சேர்ந்து நிகழ்வதற்கான நிகழ்திறன்களின் பெருக்குத் தொகையை மொத்த உறுப்புகளால் பெருக்குவதற்குச் சமமாகும்: $A_r\ B_S$ இரண்டும் சேர்ந்து நிகழ்வதற்கான நிகழ்திறன்

$$P(A_r\ B_S) = P(A_r) \cdot P(B_S)$$

$$= \frac{(A_r)}{N} \cdot \frac{(B_S)}{N}$$

இந்த மதிப்பை மொத்த உறுப்புகளால் பெருக்கும்போது $A_r\ B_s$ இரண்டும் சேர்ந்து நிகழ்வதற்கான கணக்கியல் எதிர் பார்ப்பு நிகழ்வெண் கிடைக்கிறது. இதனை $(A_r\ B_s)_e$ எனக் குறிப் பிடுவோம்.

$$(A_r\ B_s)_e = P(A_r\ B_s) \cdot N$$

$$= \frac{(A_r)}{N} \cdot \frac{(B_s)}{N} \cdot N$$

$$= \frac{(A_r)\ (B_s)}{N}$$

$A_r\ B_s$ என்ற அறைக்கான எதிர்பார்ப்பு நிகழ்வெண், $A_r\ B_s$ அமையும் நிமிர்நிலை வரிசை மொத்தத்தையும், படுகைநிலை வரிசை மொத்தத்தையும் பெருக்கிவரும் தொகையை மொத்த உறுப்புகளின் எண்ணிக்கையால் வகுப்பதற்குச் சமமாகும். எனவே, இணைப்பட்டியலில் ஓர் அறைக்கான கணக்கியல் எதிர் பார்ப்பு நிகழ்வெண்ணைக் கணக்கிட, அவ்வறை அமைந்த நிமிர். படுகைநிலை வரிசைகளின் மொத்தங்களைப் பெருக்கிவரும் மதிப்பை இணைப்பட்டியலின் மொத்த உறுப்புகளின் எண்ணிக் கையால் வகுக்க வேண்டும்.

சோதனை முடிவுப்படி கிடைக்கும் மதிப்புகளை O_1 O_2.........O_r எனவும், கணக்கியல் முடிவுகளை e_1, e_2.........e_n எனவும் குறிப்பிடுவோம். இணைப்பட்டியலில் உள்ள எல்லா அறைகளுக்குமான கணக்கியல் எதிர்பார்ப்பு நிகழ்வெண்களைக் கணக்கிட்டால், அமையும் பட்டியலின் துணை மொத்தங்கள் (row, column totals) கொடுக்கப்பட்ட பட்டியலுக்கான மொத்தங்களுக்குச் சமமாக இருக்கும். இக் கட்டளைப்படி

$$\sum_{i=1}^{n} (O_i - e_i) = 0 \text{ என அமையும்.}$$

ஓர் இணைப்பட்டியலுக்கான சோதனை முடிவுகளிலிருந்து கணக்கியலாக எதிர்பார்க்கும் முடிவுகளைக் கழித்துவரும் வித்தியாசங்களின் கூட்டுத்தொகை பூச்சியத்திற்குச் சமமாகும் என்பதை இது சுட்டுகிறது.

கை-வர்க்கம் (Chi-square) எனும் அளவையை

$$\chi^2 = \sum_{1}^{n} \frac{(O_i - e_i)^2}{e_i} \text{ என வரையறுக்கிறோம்.}$$

ஒவ்வோர் அறைக்குமான சோதனை முடிவிலிருந்து கணக்கியலாக எதிர்பார்க்கும் முடிவுகளைக் கழித்துவரும் வித்தியாசங்களின் வர்க்கத்தை, கணக்கியலாக எதிர்பார்க்கும் நிகழ்வெண்ணால் வகுக்கக் கிடைக்கும் மதிப்புகளின் கூட்டுத்தொகைக்கு இது சமமாகும். இதன் கூட்டுச் சராசரியே 'சராசரி வர்க்க இணைப்பு' எனப்படும். இவ் விணைப்பைக் குறிப்பிடவே கார்ல் பியர்சன் (Karl Pearson) ஓர் இணைப்புக் கெழுவை அமைத்துள்ளார். இதனை

$$C = \sqrt{\frac{\chi^2}{N+\chi^2}} = \sqrt{\frac{\phi^2}{1 + \phi^2}}$$

எனக் கொள்கிறோம். (ϕ^2 என்பது, $\frac{\chi^2}{N}$ ஐக் குறிக்கும்.) இதுவே ' இணைப்புக் கெழு ' (co-efficient of contingency) எனப்படும்.

பொதுவாக இதன் மதிப்பிற்கு எந்தக் குறியும் (+, −) கொடுக்கப்படுவதில்லை. இதன் மதிப்பு இரு பண்புகளின் இயைபுடைய அல்லது இயைபற்ற தன்மையை நமக்குத் தெளிவுபடுத்துவதைக் குறிக்கும். இதன் மதிப்பு, ϕ^2-ன் மதிப்பு அதிகப்பட அதிகப்பட ஒன்று என்ற நிலையை அடைந்தாலும், இதன் அதிகப்படியான மதிப்பு, பண்புகள் பிரிக்கப்படும் அளவுகளுக்குத் தக்கவாறு அமையும். அதாவது இதன் மதிப்பு m, n (number of rows and columns) ஆகியவற்றின் மதிப்பைச் சார்ந்தமையும்.

காட்டு :

சகோதரர், சகோதரிக்களுக்கான குணங்களின் பகுப்பிற்கான இயைபைக் கணக்கிடுவோம்.

சகோதரிகளின் குணம் / சகோதரர்களின் குணம்	முன் கோபி	இயற் பண்பு	சிடுசிடுப்பு	
முன் கோபி	850 O_1	571 O_2	580 O_3	2001
இயற்பண்பு	618 O_4	593 O_5	455 O_6	1666
சிடுசிடுப்பு	540 O_7	456 O_8	457 O_9	1453
	2008	1620	1492	5120

இணைப்பட்டியலில் உள்ள ஒவ்வோர் அறைக்குமான எதிர் பார்ப்பு நிகழ்வெண்ணைக் கணக்கிடுவோம். இவை முறையே $e_1, e_2 \ldots\ldots e_8, e_9$ எனக் குறிக்கப்பட்டுள்ளன.

$$e_1 = \frac{2001 \times 2008}{5120} = 785$$

$$e_2 = \frac{2001 \times 1620}{5120} = 633$$

$$e_3 = \frac{2001 \times 1492}{5120} = 583$$

$$e_4 = \frac{1666 \times 2008}{5120} = 653$$

$$e_5 = \frac{1666 \times 1620}{5120} = 527$$

$$e_6 = \frac{1666 \times 1492}{5120} = 486$$

$$e_7 = \frac{1453 \times 2008}{5120} = 573$$

$$e_8 = \frac{1453 \times 1620}{5120} = 460$$

$$e_9 = \frac{1453 \times 1492}{5120} = 423$$

கணக்கியல் எதிர்பார்ப்பிற்கான இணைப்பட்டியல்

சகோதரிகளின் குணம் / சகோதரர்களின் குணம்	முன்கோபி	இயற்பண்பு	சிடுசிடுப்பு	மொத்தம்
முன்கோபி	785 e_1	633 e_2	583 e_3	2001
இயற்பண்பு	653 e_4	527 e_5	486 e_6	1666
சிடுசிடுப்பு	573 e_7	460 e_8	423 e_9	1453
மொத்தம்	2008	1620	1492	5120

$$\chi^2 = \frac{(O_1 - e_1)^2}{e_1} + \frac{(O_2 - e_2)^2}{e_2} + \ldots\ldots + \left(\frac{O_9 - e_9}{e_9}\right)^2$$

$$= \frac{(850 - 785)^2}{785} + \frac{(571 - 633)^2}{633} + \ldots\ldots + \frac{(457 - 423)^2}{423}$$

$$= \frac{4225}{785} + \frac{3844}{633} + \frac{9}{583} + \frac{1225}{653} + \frac{4356}{527} + \frac{961}{480}$$

$$+ \frac{900}{570} + \frac{16}{460} + \frac{1156}{429}$$

$$= 5{\cdot}38 + 6{\cdot}07 + 0{\cdot}02 + 1{\cdot}88 + 8{\cdot}27 + 1{\cdot}98 + 1{\cdot}58$$

$$+ 0{\cdot}03 + 2{\cdot}73$$

$$\therefore \quad \chi^2 = 27{\cdot}94$$

$$C = \sqrt{\frac{\chi^2}{N + \chi^2}}$$

$$= \sqrt{\frac{27{\cdot}94}{5120 + 27{\cdot}94}}$$

$$C = \sqrt{{\cdot}0054} = 0{\cdot}074$$

பயிற்சி :

1. குழந்தைகளின் வளர்ச்சியைக் காணச் செய்யப்பட்ட சோதனை முடிவுகள் பின்வருமாறு :

A = குறைபாடுடைய தோற்றம்.

B = இயல்பான நிலை.

C = குறை உணவு.

$N = 1000$	$(AB) = 338$
$(A) = 877$	$(AC) = 143$
$(B) = 1086$	$(BC) = 135$
$(C) = 286$	$(ABC) = 57$

இப் பகுப்புமுறையின் மற்றப் பண்புகளின் கூட்டிற்கான நிகழ்வெண்களைக் கணக்கிடு.

2. கொடுக்கப்பட்ட உயர்படி நிகழ்வெண்களிலிருந்து (Ultimate frequencies) மொத்த உறுப்புகளையும், நேர், எதிர் பிரிவுகளுக்கான நிகழ்வெண்களையும் கணக்கிடு.

$(AB) = 733$	$(A\beta) = 840$
$(\alpha B) = 699$	$(\alpha\beta) = 783$

3. கொடுக்கப்பட்ட முடிவுகளிலிருந்து நேர்பண்புகளின் நிகழ்வெண்களைக் கணக்கிடு.

$(ABC) = 149$	$(\alpha BC) = 204$
$(AB\gamma) = 738$	$(\alpha B\gamma) = 1762$
$(A\beta C) = 225$	$(\alpha\beta C) = 171$
$(A\beta\gamma) = 1195$	$(\alpha\beta\gamma) = 2842$

4. ‘இயைபு கெழுவினை’ வரையறுத்து அதற்கான மதிப்புப்பற்றி விளக்கு. படிப்பிற்கும் வேலையில்லாத் தன்மைக்குமான (கிராமத்திலும் நகரத்திலும்) இயைபுபற்றிக் கணக்கிடச் சேகரிக்கப்பட்ட முடிவுகளிலிருந்து அதன் இயைபு கெழுவினைக் கணக்கிட்டு அவற்றில் உள்ள வித்தியாசத்தை விளக்குக.

	கிராமம்	நகரம்
மொத்த நபர்கள்	25 லட்சம்	200 லட்சம்
படித்தவர்கள்	10 ,,	40 ,,
வேலையில்லாதவர்கள்	5 ,,	4 ,,
படித்தவர்களில் வேலையில்லாதவர்கள்	3 ,,	4 ,,

5. முறைக்காய்ச்சலைத் தடுக்கச் செய்யப்பட்ட சோதனையில் 4000 பேர் உள்ள கிராமத்தில் 960 பேருக்குக் கொயினா மருந்து கொடுக்கப்பட்டது. சோதனை முடிவுகள் 2 × 2 இணைப் பட்டியலில் குறிக்கப்பட்டுள்ளன.

	முறைக்காய்ச்சல் உள்ளவர்கள்	முறைக்காய்ச்சல் இல்லாதவர்கள்	
கொயினா மருந்து சாப்பிட்டவர்கள்	40	920	960
கொயினா மருந்து சாப்பிடாதவர்கள்	540	2500	3040
	580	3420	4000

மருந்து சாப்பிட்டதற்கும் முறைக்காய்ச்சல் வராமலிருப்பதற்கும் உள்ள இயைபைக் (association) காண்.

6. மாடுகளுக்குக் காசநோய் (Tuberculosis) வராமலிருக்கத் தடுப்பூசி போட்டதில் பின்வரும் முடிவுகள் குறிக்கப்பட்டுள்ளன.

	இறந்தவை அல்லது பாதிக் கப்பட்டவைக்	பாதிக்கப் படாதவை	
ஊசி போடப் பட்டவை	12	26	38
ஊசி போடப் படாதவை	16	12	28
	28	38	66

ஊசி போடுவதற்கும், வியாதியால் பாதிக்கப்படாமலிருப்பதற்கும் உள்ள இயைபைக் காண்.

7. 1000 மணத் தம்பதிகளுக்கான அளவுகளில் கிடைத்த முடிவுகள் பின்வருமாறு. 800 கணவன்மார்கள் மனைவிகளின் ஓர் அளவினைக் காட்டிலும் அதிகமான அளவைப் பெற்றவர்களாகவும், 700 கணவன்மார்கள் மனைவிகளின் மற்றோர் அளவினைக் காட்டிலும் அதிகமான அளவினைப் பெற்றவர்களாகவும், 420 பேர் இரண்டு அளவிலும் அதிகமான அளவினைப் பெற்றவர்களாகவும் குறிப்பிடப்படுகிறது. இதில் உள்ள இசைவற்ற தன்மையை நிரூபி.

8. கொடுக்கப்பட்டுள்ள சோதனை முடிவுகள் இசைவுடையனவா என்பதைச் சரிபார்.

$N = 600$ $\qquad (\lambda B) = 400$

$(A) = 250$ $\qquad (A\beta) = 200$

9. A என்பது உரமிட்ட வயல்களையும், B என்பது ஆற்றுப் பாசன வயல்களையும், C என்பது தரமான விதைகள் விதைக்கப்பட்ட வயல்களையும் குறிக்குமிடத்து அடியிற்காணும் முடிவுகள் பெறப்படுகின்றன.

$N = 1000$ $\qquad (AB) = 189$

$(A) = 510$ $\qquad (AC) = 140$

$(B) = 490$ $\qquad (BC) = 85$

$(C) = 427$

இவை ஒன்றுக்கொன்று இசைவுடையனவா என்பதைச் சரிபார்.

10. அடியிற் காணும் அளவுகள் இயைவுடையனவா என்பதைச் சோதனை செய்.

(a) $(A) = 48$, $(AB) = 50$, $(B) = 50$, $(AB) = 60$

(b) $(B) = 172$, $(AB) = 676$, $(AC) = 286$, $(BC) = 270$

11. அடியிற் காணும் அளவுகளில் A, B-க்கு உள்ள இயைபையும், C என்ற பண்பமைந்த நிலையில் A, B-க்குள்ள பகுதி இயைபின் தன்மையையும் கணக்கிடு.

$N = 100$ $(C) = 40$

$(A) = 70$ $(AB) = 44$ $(BC) = 34$

$(B) = 54$ $(AC) = 35$ $(ABC) = 32$

12. மக்கட் கணக்கெடுப்பில் பார்வையற்ற தன்மைக்கும் பித்துப்பிடித்த தன்மைக்குமான பகுப்பு அடியிற் காணும் முடிவைத் தருகிறது.

	வயதின் பிரிவு	
	15–25	75-க்குமேல்
மக்கள் தொகை	270000	160000
பார்வையற்றோர்	1000	2000
பித்துப் பிடித்தோர்	6000	1000
பார்வையற்றோரில் பித்துப் பிடித்தோர்	19	9

இரண்டு பிரிவிலும் பார்வையற்ற தன்மைக்கும் பித்துப்பிடித்த நிலைக்கும் உள்ள இயைபைக் கணக்கிடு. ஒரு பிரிவைக் காட்டிலும் மற்றொரு பிரிவில் இதற்கான இயைபு அதிகமா அல்லது இரு பிரிவிலும் சமமானதா என்பதை விளக்கு.

13. கண்களின் நிறத்திற்கும் கூந்தலின் நிறத்திற்கும் உள்ள இயைபைக் காணச் செய்யப்பட்ட சோதனை முடிவுகள் அடுத்த பக்கப் பட்டியலில் தரப்பட்டுள்ளன.

கூந்தலின் நிறம் / கண் நிறம்	செவ்விய நிறம் (Fair)	காவி நிறம் (Brown)	கறுப்பு (Black)	சிவப்பு (Red)	
நீலம் (Blue)	1768	807	189	47	2811
சாம்பல் நிறம் (Grey)	946	1387	746	53	3132
காவிநிறம் (Brown)	115	438	288	16	857
	2829	2632	1223	116	6800

14. அடியிற் காணும் அளவுகள் இசைவுடையனவா என்பதைச் சரிபார்.

$$N = 1000 \qquad (AB) = 44$$

$$(A) = 525 \qquad (AC) = 35$$

$$(B) = 312 \qquad (BC) = 34$$

$$(C) = 470 \qquad (ABC) = 32$$

15. கண்பார்வைக்கும், கைப்பழக்கத்திற்கும் உள்ள இயைபைக் காணச் செய்யப்பட்ட சோதனை முடிவுகள் பின் வருமாறு:

கண் / கை	இடக் கண் பழக்கம்	இயல் நிலைப் பரவல்	வலக் கண் பழக்கம்	மொத்தம்
இடக்கைப் பழக்கம்	32	60	28	120
இயல்நிலைப் பழக்கம்	25	27	18	70
வலக்கைப் பழக்கம்	56	104	50	210
மொத்தம்	113	191	96	400

கண்பார்வைக்கும் கைப்பழக்கத்திற்கும் உள்ள இயைபைக் **கணக்கிடு.**

8. உடன்தொடர்பும் தொடர்புகோடுகளும்

(Correlation and Regression)

சேர்க்கப்பட்ட அளவுகளின் போக்கு, சிதறல் முகட்டின் தன்மை, உருவமைப்பு ஆகியவற்றிற்கான புள்ளியியல் அளவை களைக் கண்டோம். உயிர்ப் புள்ளியியலில் மிகவும் அதிகமாகப் பயன்படுத்தப்படும் மற்றோர் அளவை: உடன்தொடர்பு கெழு (correlation co-efficient) ஆகும். இரண்டு மாறிகளின் அளவுகளில் உள்ள காரணத் தொடர்பை (causal relationship) உடன் தொடர்பு என்கிறோம். ஆய்விற்குரிய அளவுகளில் உள்ள தொடர்பு, இரண்டு மாறிகளின் தொடர்பின் தன்மையில் உள்ள காரண காரியத்தை (cause and effect) விளக்குவதாக இருக்க வேண்டும். உடன்தொடர்பு இரு மாறிகளும் ஒன்றையொன்று சார்ந்துள்ள (interdenpendence) அல்லது அவற்றின் கூட்டுச் சார்பினை விளக்குவதாகும். ஒரு மாறியில் ஏற்படும் மாற்றம் மற்ற மாறியில் ஏற்படும் மாற்றத்தைச் சார்ந்தமைவதையே இது குறிப்பிடும். நம் உயரத்திற்கும் எடைக்கும் உள்ள தொடர்பு, மழை அளவிற்கும் விளைபலனுக்கும் உள்ள தொடர்பு, தந்தை, அவர்தம் மக்களின் உயரங்களுக்குமுள்ள தொடர்பு முதலியவற்றை உடன்தொடர்பிற்குக் காட்டாகச் சொல்லலாம். இம் மாறிகளில் ஏதாவதொன்றில் ஏற்படும் மாற்றம் மற்றதிலும் மாற்றத்தை ஏற்படுத்தும் நிலையிலோ அல்லது இரண்டு மாறி களிலும் இயைபான மாற்றம் விளைவிக்கும் நிலையிலோ பொது வான காரணங்கள் இருப்பதை உணரலாம்.

இரு மாறிகளில் ஏற்பட்டுள்ள மாற்றம் எத்தகைய உறவை உடையதாயினும், ஒரு மாறியில் ஏற்படும் மாற்றம் மற்ற மாறியில் ஏற்பட்ட மாற்றத்தினால் உண்டானதற்கான காரணமிருந்தா லன்றி இரு மாறிகளும் தொடர்புடையன எனச் சொல்ல முடி யாது. தமிழ்நாட்டில் உற்பத்தியாகும் நெல்விளைவிற்கும், ஆந்தி ராவில் கிடைக்கும் மணிலா உற்பத்திக்கும் உள்ள தொடர்பின்

தன்மை பொருளற்றதாகும். தமிழ்நாட்டில் கிடைக்கும் அதிக நெல் விளைவிற்கான காரணங்கள், ஆந்திராவில் உற்பத்தியாகும் மணிலா விதையின் அதிகப் பலனுக்கும் காரணமாகச் சொல்லமுடியாது. இரண்டிலும் ஏற்பட்டுள்ள மாற்றங்கள் ஒன்றையொன்று சார்ந்தமையாமல் தனித்தனிக் காரணங்களுக்கானதால், இவை இரண்டும் தொடர்புடையன எனச் சொல்ல முடியாது. இத்தகைய தொடர்பு போலித் தொடர்பு (spurious correlation) எனப்படும். பல ஆண்டுகளுக்குக் கொடுக்கப்பட்ட தற்கொலை அளவுகளும், வானொலிப் பெட்டி விற்பனை அளவுகளும் ஒரு தொடர்பைக் கொடுப்பதாகத் தெரிந்தாலும், இரண்டிலும் ஏற்பட்ட மாற்றங்கள் தனித்தனிக் காரணங்களுக்கானவை என்பதை உணர முடிகிறது. வானொலிப் பெட்டி விற்பனையில் ஏற்பட்ட மாற்றத்திற்கான காரணங்கள் தற்கொலை எண்ணிக்கையில் ஏற்பட்ட மாற்றத்திற்கானவை எனச் சொல்ல முடியாது. இரண்டளவுகளிலும் தொடர்பு இருப்பதுபோல் தோன்றினாலும், ஒன்றில் ஏற்பட்ட மாற்றம் மற்றதில் ஏற்பட்ட மாற்றத்திற்கான காரணமாகக் கொள்ள முடியாதாகையால், இவற்றை உடன்தொடர்புள்ள மாறிகள் எனச் சொல்ல முடியாது. இத்தகைய தொடர்பைப் பொருளில்லாத் தொடர்பு (non-sense correlation) எனச் சொல்கிறோம். இரு மாறிகளின் தொடர்பு கணக்கிடப்படுவதற்குமுன் அவற்றில் போலி மற்றும் பொருளில்லாத் தொடர்பிற்கான காரணங்கள் உண்டா என்பதைத் தெளிவுபடுத்திக்கொள்ள வேண்டும். இரண்டு மாறிகளுக்குள்ள காரணத் தொடர்பே உண்மையான தொடர்பாகும்.

சிதறல் விளக்கப் படம் (Scatter Diagram)

உடன்தொடர்பை ஓர் அளவையால் குறிப்பிடுமுன், கொடுக்கப்பட்ட மாறிகளின் தொடர்பின் தன்மையைச் சிதறல் விளக்கப் படத்தால் குறிப்பிடும் முறையைப் பார்ப்போம். உயரத்திற்கும் எடைக்கும் உள்ள 12 நபர்களின் அளவுகள் கொடுக்கப்பட்டுள்ளன.

உயரங்களை x-அச்சிலும், எடைகளை y-அச்சிலும் சரியான அலகுகளைக் (scale) கொண்டு (60, 110), (60, 135)......(70, 160) என்ற அளவுகளுக்கான புள்ளிகள் குறிக்கப்படுகின்றன. வரைபடத்தில் உள்ள ஒவ்வொரு புள்ளியும் ஒருவர்தம் உயரத்தையும் எடையையும் குறிக்கும். இப் புள்ளிகள் அமையப்பெற்ற வரைபடமே சிதறல் விளக்கப் படமெனச் சொல்லப்படுகிறது.

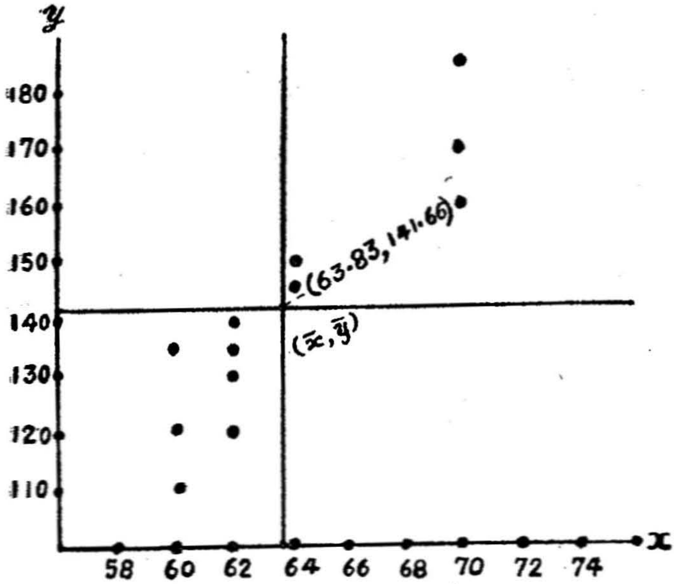
உயரம் அங். x	எடை பவு. y
60	110
60	135
60	120
62	120
62	140
62	130
62	135
64	150
64	145
70	170
70	185
70	160

சிதறல் விளக்கப் படத்திலிருந்து உடன்தொடர்பின் போக்கை அறிய, வரைபடத்தில் உள்ள புள்ளிகள் யாவும் மிக அருகில் அமையும்படியாக ஒரு நேர்கோடு வரையப்படுகிறது. இந் நேர்கோடு எவ்வாறு கணக்கியலாக வரையறுக்கப்படுகிறது என்பதைக் காண உள்ளோம். வரையப்பட்ட நேர்கோட்டைச் சுற்றிலும் நெருக்கமாக உள்ள புள்ளிகளின் தொடர்பு, மாறிகளின் நெருக்கமான தொடர்பை உணர்த்தும். நேர்கோட்டிலிருந்து சிதறி அமைந்துள்ள புள்ளிகளின் தன்மை தொடர்பின் விலக்கத்தைக் காட்டுவதாகும்.

x மாறி, y மாறி ஆகியவற்றின் கூட்டுச் சராசரியைக் கணக்கிட்டு, இவ்விரு மதிப்புகளுக்கான புள்ளியின் $(\overline{x}, \overline{y})$ வழியாக x-அச்சு, y-அச்சிற்கு இணையாக நேர்கோடுகள் வரைவோம். (63·83, 141·66) ஆகிய மதிப்புகளுக்கான புள்ளியின் வழியாக இக் கோடுகள் அமையும். சிதறல் விளக்கப் படத்தில் உள்ள புள்ளிகள் நான்கு பாகங்களாகப் பிரிக்கப்பட்டு ஒரு குறிப்பிட்ட

இடத்தைப் பெறுவதைக் காணலாம். பெரும்பாலான புள்ளிகள் முதல், மூன்றாவது பகுதிகளில் அமையும்போது இரு மாறிகளுக்கான நேர் தொடர்பையும் (positive correlation), இரண்டாவது

சிதறல் விளக்கப் படம்



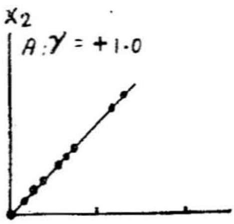
படம் 8.1

நான்காவது பகுதிகளில் அமையும்போது எதிர்த் தொடர்பையும் (negative correlation) குறிக்கும். ஒரு மாறியின் பெரிய மதிப்புகள் (large values) மற்ற மாறியின் பெரிய மதிப்புகளுடனும், ஒரு மாறியின் குறைந்த மதிப்புகள் (small values) மற்ற மாறியின் சிறிய மதிப்புகளுடனும் சேர்ந்தமையும்போது, சிதறல் விளக்கப்படப் புள்ளிகள் முதல், மூன்றாவது பகுதிகளில் குவிந்தமையும். ஒரு மாறியின் குறைந்த மதிப்புகளுடன் மற்ற மாறியின் அதிகமான மதிப்புகளும், ஒரு மாறியின் அதிகமான மதிப்புகளுடன் மற்ற மாறியின் குறைந்த மதிப்புகளும் சேர்ந்தமையும்போது சிதறல் விளக்கப்படத்தின் புள்ளிகள் இரண்டாவது, நான்காவது பகுதிகளில் குவியும். இத் தன்மையே புள்ளிகளின் இடத்திற்குத் தக்கவாறு நேர் தொடர்பு, எதிர்த் தொடர்பு எனக் கொள்ள இடமளிக்கிறது.

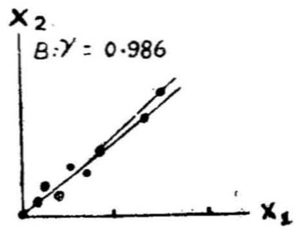
தொடர்பு கெழுவின் தன்மைபற்றித் தெரிய 9 சோடி உறுப்புகளுக்கான சிதறல் விளக்கப்படங்கள் வரையப்பட்

டுள்ளன. $n = 9, \ \bar{x} = 12, \ \bar{y} = 6, \ \Sigma x_i^2 = 576, \ \Sigma y_i^2 = 144.$ படத்திற்குப் படம் Σxy-ல் மட்டும் மாற்றம் ஏற்படும்படி அமைக்கப்பட்டுள்ளது.

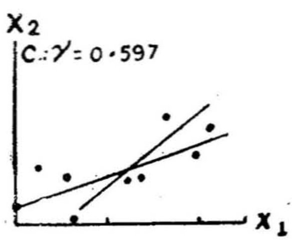
மாறுபட்ட மதிப்புகள் பெறும் உடன்தொடர்பு
கெழுக்களைக் கொடுக்கும் சிதறல் விளக்கப் படங்கள்



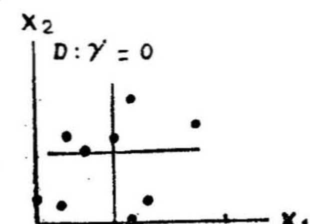
$x_1 =$ 0 4 6 8 12 14 16 22 26
$x_2 =$ 0 2 3 4 6 7 8 11 13



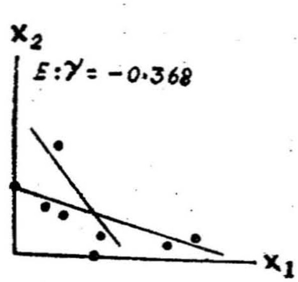
$x_1 =$ 0 4 6 8 12 14 16 22 26
$x_2 =$ 0 2 4 3 7 6 8 11 13



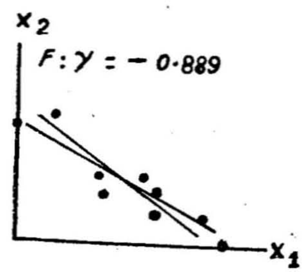
$x_1 =$ 0 4 6 8 12 14 16 22 26
$x_2 =$ 2 8 0 6 3 4 13 7 11



$x_1 =$ 0 4 6 8 12 14 16 22 26
$x_2 =$ 4 3 8 6 7 13 2 11 0



$x_1 =$ 0 4 6 8 12 14 16 22 26
$x_2 =$ 8 7 6 13 0 2 11 3 4



$x_1 =$ 0 4 6 8 12 14 16 22 26
$x_2 =$ 11 13 8 4 7 6 3 2 0

படம் 8·2

தொடர்பு கெழு (Correlation co-efficient)

சிதறல் விளக்கப்படத்தின் உதவியால் இரு மாறிகளில் உள்ள தொடர்பின் தன்மை தெரியவருகிறது. இரு மாறிகளுக்கான தொடர்பின் நெருக்கத்தைக் குறிப்பிடும் அளவையே தொடர்பு கெழு எனச் சொல்லப்படுகிறது.

$x_1\ x_2 \ldots\ldots x_n$ என்ற அளவுகளுள்ள x மாறியின் மாறுபாட்டை $\frac{1}{n}\Sigma(x_i - \bar{x})^2$ என்ற அளவையாலும், $y_1\ y_2 \ldots\ y_n$ என்ற y-மாறியின் மாறுபாட்டை $\frac{1}{n}\Sigma(y_i - \bar{y})^2$ என்ற அளவையாலும் குறிப்பிட்டோம். இரு மாறிகளிலும் சேர்ந்தாற்போல் ஏற்படும் உடன்மாறுபாட்டினை (co-variance) $\frac{1}{n}\Sigma(x_i - \bar{x})(y_i - \bar{y})$ எனக் குறிப்பிடுவோம். மாறுபாட்டளவைகள், உடன்மாறுபாட்டளவையை உபயோகப்படுத்திக் கார்ல் பியர்சனுல் (Karl Pearson) தொடர்பு கெழு வரையறுக்கப்பட்டுள்ளது.

$$\text{தொடர்புகெழு} = rxy = \frac{\frac{1}{n}\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n}\Sigma(x_i - \bar{x})^2\ \frac{1}{n}\Sigma(y_i - \bar{y})^2}}$$

தொடர்புகெழுவின் பகுதியில் உள்ள இரு அளவைகளையும் கணக்கிடும் முறையை அறிவோம். இவையிரண்டும் x, y மாறிகளுக்கான திட்ட விலக்கங்களாகும். தொகுதியில் அமைந்த உள் மாறுபாட்டளவையைக் கணக்கிடுவதும் எளிதானதாகும். மாறுபாட்டளவைகள் கணக்கிட அமைக்கப்படும் xy மாறிகளுக்கான விலக்கங்கள் இரண்டையும் பெருக்கிவரும் மதிப்புகளின் மொத்தத்தை உறுப்புகளின் மொத்த எண்ணிக்கையால் வகுத்து வரும் மதிப்பிற்கு இது சமமாகும்.

$$\mu_{11} = \frac{1}{n}\Sigma(x_i - \bar{x})(y_i - \bar{y})$$

$$S_x = \sqrt{\frac{1}{n}\Sigma(x_i - \bar{x})^2}$$

$$S_y = \sqrt{\frac{1}{n}\Sigma(y_i - \bar{y})^2}$$

$$rxy = \frac{\mu_{11}}{S_x \cdot S}$$

தொடர்பு கெழு கணக்கிட உதவும் பட்டியல்

x_i	y_i	$x_i - \bar{x}$	$y_i - y$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$\Sigma (x_i - \bar{x}) (y_i - \bar{y})$
x_1	y_1	$x_1 - x$	$y_1 - \bar{y}$	$(x_1 - \bar{x})^2$	$(y_1 - \bar{y})^2$	$(x_1 - \bar{x}) \times (y_1 - \bar{y})$
x_2	y_2	$x_2 - \bar{x}$	$y_2 - \bar{y}$	$(x_2 - x)^2$	$(y_2 - \bar{y})^2$	$(x_2 - \bar{x}) \times (y_2 - \bar{y})$
.	.	.	.	.	.	
.	.	.	.	.	.	
.	.	.	.	.	.	
.	.	.	.	.	.	
x^n	y_n	$x^n - \bar{x}$	$y_n - \bar{y}$	$(x_n - \bar{x})^2$	$(y_n - \bar{y})^2$	$(x_n - \bar{x}) (y_n - \bar{y})$
Σx_i	Σy_i			$\Sigma(x_i - \bar{x})^2$	$\Sigma(y_i - \bar{y})^2$	$\Sigma (x_i - \bar{x}) (y_i - \bar{y})$

$$\bar{x} = \frac{\Sigma x}{n} \qquad \bar{y} = \frac{\Sigma y_i}{n}$$

$$\mu_{11} = \frac{1}{n} \Sigma (x_i - \bar{x}) (y_i - \bar{y})$$

$$S_x^2 = \frac{1}{n} \Sigma (x_i - \bar{x})^2$$

$$S_y^2 = \frac{1}{n} \Sigma (y_i - \bar{y})^2$$

$$rxy = \frac{\mu_{11}}{S_x \cdot S_y}.$$

காட்டு :

அண்ணன், தங்கைகட்கான உயரங்களில் உள்ள தொடர்பினைக் கணக்கிடுவோம்.

x — அண்ணன் உயரம் (அங்.)

y — தங்கை உயரம் (அங்.)

x_i	y_i	$(x-69)$	$(y-64)$	$(x_i-69)^2$	$(y_i-64)^2$	$(x_i-69)(y_i-64)$
71	69	+2	+5	4	25	10
68	64	−1	0	1	0	0
66	65	−3	+1	9	1	−3
67	63	−2	−1	4	1	2
70	65	+1	+1	1	1	1
71	62	+2	−2	4	4	−4
70	65	+1	+1	1	1	1
73	64	+4	0	16	0	0
72	66	+3	+2	9	4	6
65	59	−4	−5	16	25	20
66	62	−3	−2	9	4	6

$$\bar{x} = \frac{759}{11} = 69 \qquad \bar{y} = \frac{704}{11} = 64$$

$$S_x = \sqrt{\frac{1}{n}\Sigma(x_i-\bar{x})^2} = \sqrt{\frac{74}{11}} = \sqrt{6{\cdot}7273}$$

$$= 2{\cdot}593$$

$$S_y = \sqrt{\frac{1}{n}\Sigma(y_i-\bar{y})^2} = \sqrt{\frac{66}{11}} = \sqrt{}$$

$$= 2{\cdot}4495$$

$$\mu_{11} = \frac{1}{n} \Sigma (_i - \overline{x} \quad - \overline{y}$$

$$= \frac{39}{11} = 3 \cdot 545$$

$$r_{xy} = \frac{\mu_{11}}{S_x \cdot S_y} = \frac{3 \cdot 545}{2 \cdot 4495 \times 3 \cdot 545} = 0 \cdot 558$$

தொடர்பு கெழு-சிறுவிளக்கம்

தொடர்பு கெழு இரு மாறிகளின் தொடர்பின் நெருக்கத்தைக் குறிக்கும் ஓர் அளவையாகும். இதன் குறி (sign) இவற்றின் உடன்மாறுபாட்டளவையின் (co-variance) குறியை ஒத்ததாய் இருக்கும். தொடர்பு கெழுவின் பகுதியில் உள்ள x-மாறியின் திட்டவிலக்கம், y-மாறியின் திட்டவிலக்கம் ஆகிய இரண்டுமே நேரெண்ணுக (positive numbers) அமைவதால் தொடர்பு கெழுவின் குறி அதன் தொகுதியில் உள்ள உடன் மாறுபாட்டளவையின் குறியை ஏற்கும். இரு மாறிகளுக்கான நேர் தொடர்பு முழுமையாக உள்ளபோது தொடர்பு கெழு '$+1$' ஆகவும், எதிர் உறவு முழுமையாக உள்ளபோது தொடர்பு கெழுவின் மதிப்பு '-1' ஆகவும், உறவற்ற தன்மையில் பூச்சியமாகவும் அமையும் $(-1 < r < +1)$. -1-லிருந்து $+1$ வரை அமையும் இக் கெழுவின் அளவிற்குத் தக்கவாறு மாறிகளின் நெருக்கம் விளக்கப்படும்.

இவ் வாய்பாட்டின் அடிப்படையாக, இரு மாறிகளுக்கும் உள்ள தொடர்பு ஒரு நேர்கோட்டில் அமையும் எனக் கொள்ளப் படுகிறது. சிதறல் விளக்கப்படத்தில் உள்ள புள்ளிகள் ஒரு நேர்கோட்டைச் சுற்றி நெருக்கமாக அமைகின்றன எனக் கொள் கிறோம். இக் கருத்துச் சில இடங்களில் பொருந்தாமல் இருப்பதை உணரமுடியும். வயதிற்கும், உழைக்கும் சக்திக்குமான அளவு களின் சிதறல் விளக்கப்படப் புள்ளிகள், ஒரு வளைகோட்டைச் (curve) சுற்றி அமையுமேயன்றி நேர்கோட்டைச் சுற்றி அமை யாது. குழந்தைப் பருவத்திலிருந்து வயது அதிகமாக அதிகமாக உழைக்கும் திறனும் அதிகமாகும். ஒரு குறிப்பிட்ட வயதிற்குப் பின், வயது அதிகமாக அதிகமாக உழைக்கும் சக்தியும் குறைந்து கொண்டே செல்லும். இத் தன்மையே சிதறல் விளக்கப்படப் புள்ளிகள் நேர்கோட்டைச் சுற்றி அமையாமல் ஒரு 'வளை கோட்டைச் சுற்றி அமைவதற்கான காரணமாகும்.

இரு மாறிகள் தொடர்பற்றுள்ளபோது, தொடர்பு கெழு பூச்சிய மதிப்பை ஏற்கும். ஆனால், தொடர்பு கெழு பூச்சியமாக அமையும்போது இரண்டு மாறிகளும் தொடர்பற்றன என்ற முடிவை உடனடியாகச் சொல்ல முடியாது. நேர்கோட்டடிப் படையில் தொடர்பில்லை எனச் சொல்லலாமே அன்றி, தொடர்பே இல்லை எனச் சொல்ல முடியாது. இம்மாதிரி இடங்களில் தொடர்பு விகிதம் (correllation ratio) கணக்கிடப்பட்டு, இரு மாறிகளுக்குமான தொடர்பின் நெருக்கம் விளக்கப்படவேண்டும்.

தொடர்பின் விகிதம் கணக்கிடப்பட்டு, இரு மாறிகளுக்கான தொடர்பின் நெருக்கம் விளக்கப்படுவதுபோல், பகுதித் தொடர்பு (partial correlation) மற்றும் பல்தரத் தொடர்பு (multiple correlation) கணக்கிடப்பட்டுத் தொடர்பின் தன்மை தெளிவுபடுவதுமுண்டு. x–மாறியின் ஒரு மதிப்பு, y-மாறியின் பல்வேறுபட்ட காரணங்களால் பாதிக்கப்படும் நிலையை இது காட்டுகிறது. நம் உயரத்தை x-மாறியாகக் கொண்டால், இதில் ஏற்படும் மாற்றத்திற்குப் பல காரணங்கள் கூறலாம். நமது உணவு, வாழ்க்கைத் தரம், பிறப்புத் தன்மை முதலியன உயரத்தில் ஏற்படும் மாற்றத்திற்கான சில முக்கியமான காரணங்களாகக் கொள்ளப்படுகின்றன. விளைபலனில் உண்டாகும் மாற்றம் விதை பொருள், விதைநிலம், உரங்களின் தன்மை, உரத்தின் அளவு, நீர் முதலிய பல்வேறு காரணங்களால் பாதிக்கப்படு வதை உணர்கிறோம். இம்மாதிரி இடங்களில் தொடர்பு கெழு கணக்கிடுவதற்கு மாறாக 'பகுதித் தொடர்பு கெழு', அல்லது பல்தரத் தொடர்பு கெழு' கணக்கிடப்பட்டு மாறிகளுக்கிடையே உள்ள தொடர்பின் நெருக்கம் விளக்கப்படுகிறது. x-மாறியில் ஏற்படும் மாற்றத்திற்கான பல்வேறுபட்ட காரணங்களின் கூட்டு விளைவைக் கணக்கிடுவது பல்தரத் தொடர்பு கெழுவாகும். பல் வேறுபட்ட காரணங்களில் ஒன்றின் விளைவைமட்டும் கணக் கிடுவது பகுதித் தொடர்பு கெழுவாகும். பகுதித் தொடர்பு கெழு கணக்கிடும்போது குறிப்பிட்ட காரணத்திற்கு அடுத்துள்ள மற்றக் காரணங்களால் x–மாறியில் குறிப்பிடத்தக்க அளவில் மாற்றம் இல்லை என்ற அடிப்படையில் செயல்படுவோம்.

தொடர்பு கெழு கணக்கிடுதல்-சுலபமுறை

தொடர்பு கெழுவிற்கான வாய்பாட்டைச் சிறிது மாற்றி யமைப்பதால் கணக்கிடும் முறையில் சுருக்கத்தை ஏற்படுத்திக் கொள்ளலாம்.

$$rxy = \frac{\frac{1}{n}\Sigma\,(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n}\Sigma\,(x_i - \bar{x})^2 . \frac{1}{n} . \Sigma\,(y_i - \bar{y})^2}}$$

$$= \frac{\frac{1}{n}\Sigma\, x_i y_i - \bar{x}\,\bar{y}}{\sqrt{\left(\frac{1}{n}\Sigma\, x_i^2 - \bar{x}^2\right)\left(\frac{1}{n}\Sigma\, y_i^2 - \bar{y}^2\right)}}$$

x_i	y_i	x_i^2	y_i^2	$x_i\,y_i$
x_1	y_1	x_1^2	y_1^2	$x_1\,y_1$
x_2	y_2	x_2^2	y_2^2	$x_2\,y_2$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
x_n	y_n	x_n^2	y_n^2	$x_n\,y_n$
$\Sigma\, x_i$	$\Sigma\, y_i$	$\Sigma\, x_i^2$	$\Sigma\, y_i^2$	$\Sigma\, x_i\, y_i$

$$\bar{x} = \frac{\Sigma x_i}{n}$$

$$\bar{y} = \frac{\Sigma y_i}{n}$$

$$S_x = \sqrt{\frac{1}{n}\Sigma x_i^2 - (\bar{x})^2}$$

$$S_y = \sqrt{\frac{1}{n}\Sigma y_i^2 - (\bar{y})^2}$$

$$\mu_{11} = \frac{1}{n}\Sigma x_i\, y_i - \bar{x}\;\bar{y}$$

$$rxy = \frac{\mu_{11}}{S_x \cdot S_y}$$

காட்டு :

அடியிற்காணும் அளவுகளுக்கான தொடர்பு கெழுவைக் கணக்கிடுவோம்.

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
39	36	1521	1296	1404
33	29	1089	841	957
30	29	900	841	870
33	40	1089	1600	1320
43	42	1849	1764	1806
31	29	961	841	899
38	40	1444	1600	1520
34	31	1156	961	1054
36	38	1296	1444	1368
34	32	1156	1024	1088
351	346	12461	12212	12286

$$\bar{x} = \frac{351}{10} = 35 \cdot 1. \quad \bar{y} = \frac{346}{10} = 34 \cdot 6$$

$$S_x = \sqrt{\frac{12461}{10} - (35 \cdot 1)^2}$$

$$= \sqrt{1246 \cdot 1 - 1232 \cdot 01}$$

$$= \sqrt{14 \cdot 09} = 3 \cdot 754$$

$$S_y = \sqrt{\frac{12212}{10} - (34.6)^2}$$

$$= \sqrt{1221 \cdot 2 - 1197 \cdot 16}$$

$$= \sqrt{24 \cdot 04} = 4 \cdot 904$$

$$\mu_{11} = \frac{12286}{10} - (35 \cdot 1)(34 \cdot 6)$$

$$= 1228 \cdot 6 - 1214 \cdot 46$$

$$= 14 \cdot 14$$

$$rxy = \frac{\mu_{11}}{S_x \cdot S_y}$$

$$= \frac{14 \cdot 14}{(3 \cdot 754)(4 \cdot 904)} = \frac{14 \cdot 14}{18 \cdot 4096}$$

$$= 0 \cdot 768$$

$\bar{x}$, $\bar{y}$ முழு எண்களாக அமையாமலும், கொடுக்கப்பட்ட அளவுகளிலிருந்து நேரடியாகக் கணக்கிடுவது கடினமாகப்படினும், x, y மதிப்புகளுக்குத் தோராயமான இரு மதிப்புகளை எடுத்துக் கணக்கிடும் முறையில் மேலும் சுருக்கத்தை ஏற்படுத்திக் கொள்ளலாம். 'A' என்ற தோராய மதிப்பை x-மாறிக்கும், 'B'-என்ற தோராய மதிப்பை y-மாறிக்கும் கொள்வோம். x-மாறியின் மதிப்புகளிலிருந்து 'A' என்ற மதிப்பைக் கழித்தும், y-மாறியின் மதிப்புகளிலிருந்து 'B'-என்ற மதிப்புகளைக் கழித்தும் இரு புதிய மாறிகளை அமைக்கிறோம். இப்படி மாற்றியமைக்கப்பட்ட மாறியின் அளவுகள் கணக்கிடும் முறையில் எளிமையை உண்டாக்கும் படியாக A, B-ன் மதிப்புகளைத் தேர்ந்தெடுத்துக் கொள்ள வேண்டும்.

x_i	y_i	$(x_i - A)$ u_i	$(y_i - B)$ v_i	u_i^2	v_i^2	$u_i v_i$
x_1	y_1	$(x_1 - A) = u_1$	$(y_1 - B) = v_1$	u_1^2	v_1^2	$u_1 v_1$
x_2	y_2	$(x_2 - A) = u_2$	$(y_2 - B) = v_2$	u_2^2	v_2^2	$u_2 v_2$
.	.		.	.	.	.
.	.		.	.	.	.
.	.		.	.	.	.
x_n	y_n	$(x_n - A) = u_n$	$(y_n - B) = v_n$	u_n^2	v_n^2	$u_n v_n$
		Σu_i	Σv_i	Σu_i^2	Σv_i^2	$\Sigma u_i v_i$

$$\bar{u} = \frac{\Sigma u_i}{n}. \qquad \bar{v} = \frac{\Sigma v}{n}.$$

$$S_u = \sqrt{\frac{1}{n}\Sigma u_i^2 - (\bar{u})^2}$$

$$S_v = \sqrt{\frac{1}{n}\Sigma v_i^2 - (\bar{v})^2}$$

$$\mu_{11} = \frac{1}{n}\Sigma u_i v_i - (\bar{u})(\bar{v})$$

$$r_{uv} = \frac{\mu_{11}}{S_u \cdot S_v}$$

தோராய மதிப்புகளைக் கழித்து அமைக்கப்படும் புதிய மாறிகளின் தொடர்பு கெழு கொடுக்கப்பட்ட மாறிகளுக்கான தொடர்பு கெழுவிற்குச் சமமாவதைக் காணலாம். மாற்றியமைக்கப்பட்ட மாறியின் மதிப்புகளைக் கொடுக்கப்பட்ட மாறிகளுக்கு மாற்றியமைக்காமலேயே தொடர்பு கெழு கணக்கிட்டுக் கொள்ளலாம் என்பதை இது காட்டுகிறது.

$$u_i = x_i - A. \qquad v_i = y_i - B$$

$$x_i = A + u_i \qquad y_i = B + v_i$$

$$\bar{x} = A + \bar{u}. \qquad \bar{y} = B + \bar{v}.$$

$$r_{xy} = \frac{\frac{1}{n}\Sigma (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n}\Sigma (x_i - \bar{x})^2 \cdot \frac{1}{n}\Sigma (y_i - \bar{y})^2}}$$

$$= \frac{\mu_{11}}{S_x \cdot S_y}$$

$$= \frac{\frac{1}{n}\Sigma (A + u_i - A - \bar{u})(B + v_i - B - \bar{v})}{\sqrt{\frac{1}{n}\Sigma (A + u_i - A - \bar{u})^2 \ \frac{1}{n}\Sigma (B + v - B - \bar{v})}}$$

$$= \frac{\frac{1}{n}\Sigma (u_i - \bar{u})(v_i - \bar{v})}{\sqrt{\frac{1}{n}\Sigma (u_i - \bar{u})^2 \ \frac{1}{n}\Sigma (v_i - \bar{v})^2}}$$

$$= \frac{\mu_{11}}{S_u \, . \, S_v} = ruv$$

$$rxy = ruv.$$

காட்டு :

மேலே கொடுக்கப்பட்ட கணக்கில் x-மாறியின் அளவுகளுக்கு 34 ஐத் தோராய மதிப்பாகவும், $(A = 34)$, y-மாறியின் அளவுகளுக்கு 33 ஐத் தோராய மதிப்பாகவும் $(B = 33)$ கொண்டு $u_i = x_i - 34$, $v_i = y_i - 33$ என்ற புதிய மாறிகளை அமைத்து உடன்தொடர்பு கெழுவைக் கணக்கிடுவோம்.

x_i	y_i	u_i $x_i - 34$	v_i $y_i - 33$	u_i^2 $(x_i - 34)^2$	v_i^2 $(y_i - 33)^2$	$u_i \, v_i$ $(x - 34)(y_i - 33)$
39	36	+5	+3	25	9	+15
33	29	−1	−4		16	+ 4
30	29	−4	−4	16	16	+16
33	40	−1	+7	1	49	− 7
43	42	+9	+9	81	81	+81
31	29	−3	−4	9	16	+12
38	40	+4	+7	16	49	+28
34	31	0	−2	0	4	0
36	38	2	+5	4	25	10
34	32	0	−1	0	1	0
		$\Sigma u_i = +11$	$\Sigma v_i = 16$	$\Sigma u_i^2 = 156$	$\Sigma v_i^2 = 266$	$\Sigma u \, v_i = 159$

$$\bar{u} = \frac{\Sigma u_i}{n} = \frac{11}{10} = 1 \cdot 1$$

$$\bar{v} = \frac{\Sigma v_i}{n} = \frac{16}{10} = 1 \cdot 6;$$

$$S_u = \sqrt{\frac{1}{n}\Sigma u^2 - \bar{u}^2}$$

$$= \sqrt{\frac{153}{10} - (1 \cdot 1)^2}$$

$$= \sqrt{15 \cdot 3 - 1 \cdot 21} = \sqrt{14 \cdot 09}$$

$$= 3 \cdot 754$$

$$S_v = \sqrt{\frac{1}{n}\Sigma v^2 - \bar{v}^2}$$

$$= \sqrt{\frac{266}{10} - (1 \cdot 6)^2} = \sqrt{26 \cdot 6 - 2 \cdot 56}$$

$$= \sqrt{24 \cdot 04} = 4 \cdot 904$$

$$\mu_{11} = \frac{1}{n}\Sigma u\, v_i - \bar{u}\, \bar{v}_i$$

$$= \frac{159}{10} - (1 \cdot 1)(1 \cdot 6)$$

$$= (15 \cdot 9 - 1 \cdot 76) = 14 \cdot 14$$

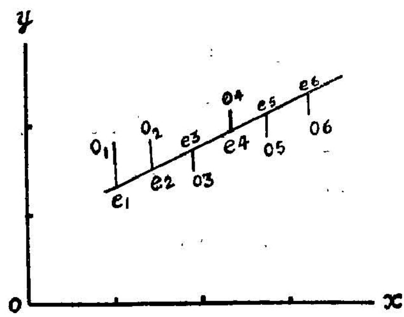
$$r_{uv} = \frac{\mu_{11}}{S_u \cdot S_v} = r_{xy}$$

$$r_{xy} = \frac{14 \cdot 4}{(3\ 754)(4 \cdot 904)} = 0\ 768$$

தொடர்பு கோடுகள் (Regression Lines)

உடன்தொடர்பு கெழு இரு மாறிகளின் தொடர்பின் நெருக்கத்தைக் காட்டும் அளவையாக உள்ளது. தொடர்பு சம்பந்தமான மற்ருேர் ஆய்வு இரு மாறிகளுக்கும் தொடர்புள்ள நிலையில், ஒரு மாறியின் குறிப்பிட்ட மதிப்பிற்கான மற்ெருரு மாறியின் மதிப்பைக் கணக்கிடுவதாகும். மழை அளவிற்கும் விளைபலனுக்கும் நேர்தொடர்பு உண்டென்பதை அறிவோம். குறிப்பிட்டவளவு மழை கிடைக்கும்போது கிடைக்கக்கூடிய விளைபலனைக் கணக்கிடுவதே நமது தற்போதையப் பிரச்சினையாகும். சிதறல் விளக்கப்படத்தில் எல்லாப் புள்ளிகளுக்கும் அருகிலமையும் நேர்கோடானது, தொடர்பின் போக்கைக் காட்டும் எனக் குறிப்பிட்டோம். கணக்கியலாக அமைக்கப்படும் இந் நேர்கோடே, ஒரு மாறியின் குறிப்பிட்ட மதிப்பிற்கான மற்ெருரு மதிப்பைக் கணக்கிட உதவுகிறது. கொடுக்கப்பட்ட மதிப்புகளிலிருந்து அமைக்கப்

படும் இந் நேர்கோட்டின்மேல் எல்லாப் புள்ளிகளும் அமைவதே சிறந்த மதிப்பீட்டைக் குறிக்கும். ஆனால், இம்மாதிரியான நேர் கோடமையப்பெறுவது அரிதாகும். கொடுக்கப்பட்டுள்ள மதிப்பு களுக்கும் நேர்கோட்டிலிருந்து கிடைக்கும் மதிப்புகளுக்கும் உள்ள வித்தியாசம் மிகவும் குறைவாக இருக்கும்படியான நேர்கோடே நன்கு பொருத்தப்பட்ட நேர்கோடாகும் (best fitted lines).



படம் 8·3

கொடுக்கப்பட்ட மதிப்புகளை $O_1, O_2 \ldots\ldots\ldots O_n$ எனவும், நேர் கோட்டிலிருந்து கிடைக்கும் மதிப்புகளை $e_1, e_2 \ldots\ldots e_n$ எனவும் குறிப்பிடுவோம். இவற்றுக்கான விலக்க வர்க்கங்களின் மொத்த மதிப்பு மிகக் குறைவாக உள்ள எதிர்பார்ப்பு அளவுகளைக் கொடுக்கும் நேர்கோடே சிறந்த நேர்கோடாகும். $\Sigma (O - e_i)^2$ மிகக் குறைவாக இருக்கவேண்டும். இம் முறையே குறைந்த வர்க்க முறை (Method of Least Squares) எனப்படும்.

x, y என்ற இரு மாறிகளில் எது சார்பு மாறி (dependent variable), எது சார்பற்ற மாறி (independent variable) என்று சொல்வது சில இடங்களில் கடினமானதாகும். இம்மாதிரி இடங் களில் x-மாறியைச் சார்பு மாறியாகக் கொண்டு ஒரு நேர்கோடும், y-மாறியைச் சார்பு மாறியாகக் கொண்டு ஒரு நேர்கோடுமாக இரு நேர்கோடுகள் வரைய வேண்டும். இந் நேர்கோடுகளே 'மாறியின் தொடர்பு கோடுகள்' (Regression lines) எனப் படும்.

$y = mx + c, x = m'y + c'$ என்ற நேர்கோட்டு அடிப்படை யில் குறைந்த வர்க்க முறையைப் பயன்படுத்தி இரு நேர்கோடு களும் கணக்கியலாக வரையறுக்கப்படுகின்றன.

y-மாறியை x-மாறியில் குறிப்பிடும் நேர்கோடு (Regression line of y on x)

$$(y_i - \overline{y}) = \frac{\mu_{11}}{\sigma x^2} (x_i - \overline{x})$$

x-மாறியை y-மாறியில் குறிப்பிடும் நேர்கோடு (Regression line of x on y)

$$(x_i - \overline{x}) = \frac{\mu_{11}}{\sigma y^2} (y_i - \overline{y})$$

$b_1 = \frac{\mu_{11}}{\sigma x^2}$, $b_2 = \frac{\mu_{11}}{\sigma y^2}$ என்பன 'மாறிகளின் தொடர்பு கெழுக்கள்' (Regression co-efficients) எனச் சொல்லப்படும். இவையிரண்டின் பெருக்கற் சராசரி (Geometric mean) தொடர்பு கெழுவிற்குச் சமமாகும்.

$$r_{xy} = \frac{\mu_{11}}{\sigma x \cdot \sigma y} = \sqrt{\frac{\mu_{11}}{\sigma x^2} \cdot \frac{\mu_{11}}{\sigma y^2}}$$

மாறிகளின் தொடர்பு கெழுக்களைக் (b_1, b_2) கணக்கிடாமல் தொடர்பு கெழுவின் உதவியாலும் நேர்கோடுகள் கணக்கிட முடியும். இதற்குத் தகுந்தாற்போல் நேர்கோடுகளின் வாய்பாடுகள் மாற்றியமைக்கப்படவேண்டும்.

y-மாறியை x-மாறியில் குறிப்பிடும் நேர்கோடு

$$(y_i - \overline{y}) = \frac{\mu_{11}}{\sigma x^2} (x_i - \overline{x})$$

$$= r \cdot \frac{\sigma y}{\sigma x} (x_i - \overline{x})$$

$$\left(\frac{y_i - \overline{y}}{\sigma y}\right) = r \left(\frac{x_i - \overline{x}}{\sigma x}\right)$$

x-மாறியை y-மாறியில் குறிப்பிடும் நேர்கோடு

$$(x_i - \overline{x}) = \frac{\mu_{11}}{\sigma y^2} (y - \overline{y})$$

$$= r \cdot \frac{\sigma x}{\sigma y} (y_i - \overline{y})$$

$$\left(\frac{x_i - \overline{x}}{\sigma x}\right) = r \cdot \left(\frac{y_i - \overline{y}}{\sigma y}\right)$$

தொடர்பு கெழுவினைக் கணக்கிட அமைத்த பட்டியலைக் கொண்டே, தொடர்பு கோடுகளுக்கான மதிப்புகளும் கணக்கிடப்படும். இதற்கெனத் தனியாக எந்த மதிப்பும் புதிதாகக் கணக்கிடத் தேவையில்லை. மாறிகளை மாற்றியமைக்கும்போது, தொடர்பு கெழுவில் மாற்றம் ஏற்படுவதில்லை என்பதைக் கண்டோம். தொடர்பு கோடுகளிலும் இத்தன்மையே இருப்பினும், கோடுகளை இறுதியாகக் குறிப்பிடும்போது கொடுக்கப்பட்ட மாறியின் அளவுகளிலேயே கொடுக்க வேண்டும்.

காட்டு :

12 சோடி கணவன் மனைவி வயதிற்கான அளவுகளின் தொடர்பு கோடுகளை அமைப்போம்.

x_i	y_i	u $(x_i-\bar{x})$	v $(y_i-\bar{y})$	u^2 $(x_i-\bar{x})^2$	v^2 $(y_i-\bar{y})^2$	uv $(x_i-\bar{x})\times(y_i-\bar{y})$
25	18	0	1	0	1	0
22	15	− 3	− 2	9	4	6
28	20	+ 3	+ 3	9	9	9
26	17	+ 1	0	1	0	0
35	22	+ 10	+ 5	100	25	50
20	14	− 5	− 3	25	9	15
22	16	− 3	− 1	9	1	3
40	21	+ 15	4	225	16	20
20	15	− 5	− 2	25	4	10
18	14	− 7	− 3	49	9	21
19	15	− 6	− 2	36	4	12
25	17	0	0	0	0	0
300	204			488	82	146

$$\bar{x} = \frac{\Sigma_i x_i}{n} = \frac{300}{12} = 25$$

$$\bar{y} = \frac{\Sigma y_i}{n} = \frac{204}{12} = 17$$

$$Sx = \sqrt{\frac{\Sigma_i (x_i - \bar{x})^2}{n}} = \sqrt{\frac{488}{12}} = \sqrt{40{\cdot}666}$$

$$Sy = \sqrt{\frac{\Sigma (y_i - \bar{y})^2}{n}} = \sqrt{\frac{82}{12}} = \sqrt{6{\cdot}8333}$$

$$\mu_{11} = \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{n} = \frac{146}{12} = 12{\cdot}166$$

y-மாறியை x-மாறியில் குறிப்பிடும் நேர்கோடு

$$(y_i - y) = \frac{\mu_{11}}{\sigma x^2} (x_i - \bar{x})$$

$$(y_i - 17) = \frac{12{\cdot}166}{40{\cdot}666} (x_i - 25)$$

$$(y_i - 17) = 0{\cdot}2992 \, (x_i - 25)$$

$$y_i = 0{\cdot}2992 \, x_i + 17 - 7{\cdot}48$$

$$y_i = 0{\cdot}2992 \, x_i + 9{\cdot}52$$

x-மாறியை y-மாறியில் குறிப்பிடும் நேர்கோடு

$$(x_i - \bar{x}) = \frac{\mu_{11}}{\sigma y^2} (y_i - \bar{y})$$

$$(x_i - 25) = \frac{12{\cdot}166}{6{\cdot}8333} (y_i - 17)$$

$$= 1{\cdot}78 \, (y_i - 17)$$

$$x_i = 1{\cdot}78 \, y_i + 25 - 30{\cdot}26$$

$$x_i = 1{\cdot}78 \, y_i - 5{\cdot}26$$

காட்டு :

அடியிற்காணும் அளவுகளுக்கான தொடர்பு கோடுகளைக் கணக்கிடுவோம். 10 மாணவர்கள் இரண்டு பாடங்களில் வாங்கிய மதிப்பெண்கள் (50-க்கு) தரப்பட்டுள்ளன.

x = கணக்கு

y = உயிரியல்

$A = 33$, $B = 34$ என்ற தோராய மதிப்புகளைக்கொண்டு u, v என்ற மாறிகளை அமைத்துத் தொடர்பு கோடுகளுக்கான மதிப்புகளைக் கணக்கிடுகிறோம்.

x_i	y_i	$x_i - 33$ u_i^2	$y_i - 34$ v_i	u_1^2	v_i^2	u_iv_i
28	30	−5	−4	25	16	20
34	34	+1	0	1	0	0
32	30	−1	−4	1	16	4
41	37	+8	+3	64	9	24
38	40	+5	+6	25	36	30
36	42	+3	+8	9	64	24
37	40	+4	+6	16	36	24
33	36	0	+2	0	4	0
32	31	−1	−3	1	9	3
33	36	0	+2	0	4	0
		+14	16	142	194	129

$\bar{u} = 1{\cdot}4$. $A = 33$. $\bar{x} = A + \bar{u} = 34{\cdot}4$.

$\bar{v} = 1{\cdot}6$. $B = 34$. $\bar{y} = B + \bar{v} = 35{\cdot}6$.

$$S_x^2 = \frac{1}{n}\Sigma u_i^2 - u^2 = \frac{142}{10} - (1{\cdot}4)^2$$

$$= 14{\cdot}2 - 1{\cdot}96$$

$$= 12{\cdot}24$$

$$= S_x^2$$

$$S_v^2 = \frac{1}{n} \Sigma v_i^2 - \bar{v}^2 = \frac{194}{10} - (1 \cdot 6)^2$$

$$= 19 \cdot 40 - 2 \cdot 56 = 16 \cdot 84$$

$$= S_y^2$$

$$\mu_{11} = \frac{1}{n} \Sigma u_i v_i - \bar{u}\ \bar{v}$$

$$= \frac{129}{10} - (1 \cdot 4)(1 \cdot 6)$$

$$= 12 \cdot 9 - 2 \cdot 24 = 10 \cdot 66$$

y-மாறியை x-மாறியில் குறிப்பிடும் நேர்கோடு

$$(y_i - \bar{y}) = \frac{\mu_{11}}{S_x^2} (x_i - \bar{x}).$$

$$(y_i - 35 \cdot 6) = \frac{10 \cdot 66}{12 \cdot 24} (x_i - 34 \cdot 4)$$

$$= 0 \cdot 871 (x_i - 34 \cdot 4)$$

$$y_i = 0 \cdot 871\, x_i + 35 \cdot 6 - 29 \cdot 9624$$

$$y_i = 0 \cdot 871\, x_i + 5 \cdot 6376$$

x-மாறியை y-மாறியில் குறிப்பிடும் நேர்கோடு

$$(x_i - \bar{x}) = \frac{\mu_{11}}{S_y^2}$$

$$(x_i - 34 \cdot 4) = \frac{10 \cdot 66}{16 \cdot 84} (y_i - 35 \cdot 6)$$

$$= 0 \cdot 633 (y_i - 35 \cdot 6)$$

$$x_i = 0 \cdot 633\, y_i + 34 \cdot 4 - 22 \cdot 5348$$

$$x_i = 0 \cdot 633\, y + 11 \cdot 8652.$$

காட்டு :

கணவன் வயது x_i	மனைவி வயது y_i	x_i^2	y_i^2	$x_i y_i$
18	17	324	289	306
19	17	361	289	323
20	18	400	324	360
21	18	441	324	378
22	18	484	324	396
23	19	529	361	437
24	19	576	361	456
25	20	625	400	500
26	21	676	441	546
27	22	729	484	594
225	189	5145	3597	4296

$$\bar{x} = 22\cdot5 \qquad \bar{y} = 1\cdot89$$

$$S = \sqrt{514\cdot5 - (22\cdot5)^2} = \sqrt{514\cdot5 - 506\cdot25}$$

$$S_x^2 = 8\cdot25$$

$$S_y = \sqrt{359\cdot7 - (18\cdot9)^2} = \sqrt{359\cdot7 - 357\cdot21}$$

$$S_y^2 = 2\cdot49$$

$$\mu_{11} = 429\cdot6 - (22\cdot5)(18\cdot9)$$

$$= 429\cdot6 - 425\cdot25$$

$$= 4\cdot35$$

x-மாறியை y-மாறியல் குறிப்பிடும் நேர்கோடு

$$(x_i - \bar{x}) = \frac{\mu_{11}}{S_y^2}(y_i - \bar{y})$$

$$(x_i - 22\cdot5) = \frac{4\cdot35}{2\cdot49}(y_i - 18\cdot9)$$

$$x_i = 1\cdot747\, y_i + 22\cdot5 - 33\cdot02$$

$$x_i = 71\cdot47 y_i - 10\cdot52$$

y-மாறியை x-மாறியில் குறிப்பிடும் நேர்கோடு

$$(y_i - \bar{y}) = \frac{\mu_{11}}{\sigma^2 x} (x_i - \bar{x})$$

$$(y_i - 18{\cdot}9) = \frac{4{\cdot}35}{8{\cdot}25} (x_i - 22{\cdot}5)$$

$$y_i = 0{\cdot}523\ x - 11{\cdot}77 + 18{\cdot}9$$

$$y = 0{\cdot}523\ x_i + 7{\cdot}13$$

தொடர்பு பட்டியல் (Correlation table)

x, y மாறிகளின் மதிப்புகள் சிறியனவாகவும் எண்ணிக்கையில் குறைந்ததாகவும் உள்ள இடங்களில் நேரிடையாகவோ அல்லது சுலப முறையைப் பயன்படுத்தியோ உடன்தொடர்பு கெழு கணக்கிடும் முறையைப் பார்த்தோம். x, y மாறிகளின் எண்ணிக்கை அதிகமாக உள்ள இடத்து, நேரிடையாகக் கணக்கிடுவதில் உள்ள சிக்கலைத் தவிர்க்க உடன்தொடர்பு பட்டியலை அமைத்துக் கணக்கிடுகிறோம். 100 கணவன்-மனைவி வயதிற்காக அமைக்கப்பட்ட தொடர்பு பட்டியலைப் பார்ப்போம்.

கணவன் வயது x / மனைவி வயது y	20–30	30–40	40–50	50–60	60–70	மொத்தம்
15–25	5	9	3			17
25–35		10	25	2		37
35–45		1	12	2		15
45–55			4	16	5	25
55–65				4	2	6
மொத்தம்	5	20	44	24	7	100

100 கணவர்களின் வயதை x-மாறியாகவும், அவர்தம் மனைவியரின் வயதை y-மாறியாகவும் கொண்டு இப் பட்டியல் அமைக்கப்பட்டுள்ளது. கணவர்களின் வயது 20–30, 30–40 40–50, 50–60, 60–70 என்ற ஐந்து இடைவெளிப் பிரிவுகளாகவும், மனைவியரின் வயது 15–25, 25–35.........55–65 என்ற ஐந்து இடைவெளிப் பிரிவுகளாகவும் பிரிக்கப்பட்டுள்ளன. 100 சோடி அளவுகளும், சரியான அறைகளில் சரிபார்க்கும் குறியால் (talley mark) அமைத்து, ஒவ்வோர் அறையிலும் அமைந்த நிகழ்வெண்களின் மதிப்புத் தரப்பட்டுள்ளது. 40–50 என்ற இடைவெளியில் அமையும் 44 கணவர்களின் வயதிற்கான மனைவியரின் வயது 15-25-ல் மூன்றும், 25–35-ல் இருபத்தைந்தும், 35–45-ல் பன்னிரண்டும், 45–55-ல் நான்குமாக 44 மனைவியரின் வயது அமைந்துள்ளது. இதேபோல் 35–45 வயதுள்ள 15 மனைவிமார்களின் கணவர் வயது 30–40-ல் ஒன்றும், 40–50-ல் பன்னிரண்டும், 50-60-ல் இரண்டுமாக 15 பேர் வயது அமைந்துள்ளது. இரண்டு மாறிகளின் தொடர்பைக் காண அமைக்கப்படும் இப் பட்டியல் தொடர்பு பட்டியல் (Correlation table) அல்லது இரட்டைமாறி நிகழ்வெண் பட்டியல் (Bivariate frequency table) எனச் சொல்லப்படும்.

உடன்தொடர்பு கெழுவைக் கணக்கிட உதவும் மதிப்புகளைத் தொடர்பு பட்டியலில் அமைக்கும் முறையைப் பார்ப்போம். தொடர்பு பட்டியலில் x, y மாறிகளுக்குக் கிடைக்கும் இரு நிகழ்வெண் பரவல்களும், தொடர்பு கெழுவின் பகுதியில் அமையும் x, y மாறிகளுக்கான திட்ட விலக்கங்களைக் கணக்கிட உதவுகின்றன. தொடர்பு பட்டியலில் அமைந்துள்ள இரு நிகழ்வெண் பரவல்களும் பின்வருமாறு:

x_i கணவன் வயது	f_i நிகழ் வெண்
20–30	5
30–40	20
40–50	44
50–60	24
60–70	7
மொத்தம்	100

y_i மனைவி வயது	f_i நிகழ் வெண்
15–25	17
25–35	37
35–45	15
45–55	25
55–65	6
மொத்தம்	100

தொடர்பு கெழுவின் தொகுதியில் உள்ள உடன்மாறுபாட்டளவையைக் கணக்கிடும் முறையைப் பார்ப்போம். நிகழ்வெண் பரவலில் ஒவ்வோர் இடைவெளிப் பிரிவிலும் அமையும் நிகழ்வெண்கள் சமமாகப் பரவி இருக்கும் என்ற அடிப்படையில் புள்ளியியல் அளவைகள் கணக்கிடப்படுவதைக் குறிப்பிட்டோம். கணவர் வயதிற்காகப் பிரிக்கப்பட்ட ஐந்து இடைவெளிப் பிரிவுகள், மனைவி வயதிற்காகப் பிரிக்கப்பட்ட ஐந்து இடைவெளிப் பிரிவுகள் கொடுக்கும் 25 சிற்றறைகளில் அமையும் நிகழ்வெண்கள் யாவும் அந்தந்த இடைவெளிப் பிரிவில் சமமாகப் பரவி இருக்கும் என்ற அடிப்படையில் உடன்மாறுபாட்டிற்கான மதிப்புகள் கணக்கிடப்படுகின்றன. ஒவ்வோர் அறையும் மூன்று அளவுகளால் சுட்டப்படும். x-மாறியின் இடைவெளிப் பிரிவு அல்லது பிரிவெண், y மாறியின் இடைவெளிப் பிரிவு அல்லது இடைவெளிப் பிரிவெண் ஆகிய இரண்டளவுகளுக்கான பொது அறையில் அமையும் நிகழ்வெண் ஆகிய மூன்றினையும் இது குறிக்கும். ஓர் அறைக்கான இம் மூன்று மதிப்புகளையும் கொண்டு, அவ்வறைக்கான உடன்மாறுபாட்டளவையின் மதிப்புக் கணக்கிடப்படுகிறது. கணவன் வயது 40–50. மனைவி வயது 35–45-ல் இருக்கும் 12 சோடி மதிப்புகளுக்கான உடன்மாறுபாட்டளவை மதிப்பைக் கணக்கிடுவோம். 40–50-க்கு 45 பிரிவெண்ணாகவும், 35-45-க்கு 40 பிரிவெண்ணாகவும் அமைகின்றன. இவ்விரண்டின் பெருக்குத் தொகையை, இவ்விடைவெளிப் பிரிவுகளின் பொது அறையில் அமையும் நிகழ்வெண்ணால் (12) பெருக்கிவரும் தொகையே இவ் வறைக்கான உடன்மாறுபாட்டளவை மதிப்பாகும். $45 \times 40 \times 12 = 21600$. இம் முறைப்படி ஒவ்வோர் அறைக்குமாகக் கணக்கிடப்பட்ட உடன்மாறுபாட்டளவைகளின் மொத்தமே தொடர்பு கெழுவின் தொகுதியில் அமையும் உடன்மாறுபாட்டளவையின் மதிப்பைக் கணக்கிட உதவுகிறது. இவ் உடன்மாறுபாட்டளவைகளின் மதிப்புகளே, கணக்கிட உதவும் தொடர்பு பட்டியலில் அறை நிகழ்வெண்களுக்கு அடியில் குறிக்கப்பட்டுள்ளது.

$$\bar{x} = \frac{4580}{100} = 45{\cdot}80$$

$$\bar{y} = \frac{3660}{100} = 36{\cdot}60$$

$$\sigma x = \sqrt{\frac{218900}{100} - (45{\cdot}8)^2}$$

$$= \sqrt{2189 - 2097{\cdot}67}$$

y \ x	20–30	30–40	40–50	50–60	60–70	மொத்தம்	பிரிவெண் y_i	$y_i \times f_i$	$y_i^2 \times f_i$	$y_i \times x_i$
15–25	5 2500	9 6300	3 2700			17	20	340	6800	11500
25–35		10 10500	25 33750	2 3300		37	30	1110	33300	47550
35–45		1 1400	12 21600	2 4400		15	40	600	24000	27400
45–55			4 9000	16 44000	5 16250	25	50	1250	62500	69250
55–65				4 13200	2 7800	6	60	360	21600	21000
மொத்தம்	5	20	44	24	7	100		3660	148200	176700
பிரிவெண் x_i	25	35	45	55	65					
$x_i \times f_i$	125	700	1980	1320	455	4580				
$x_i^2 \times f_i$	3125	24500	89100	72600	29575	218900				
$x_i \times y_i$	2500	18200	67050	64900	24050	176700				

$$= \sqrt{91 \cdot 36} = 9 \cdot 559$$

$$\sigma y = \sqrt{\frac{148200}{100} - (36 \cdot 6)^2}$$

$$= \sqrt{1482 - 1339 \cdot 56}$$

$$= \sqrt{142 \cdot 44}$$

$$= 11 \cdot 93$$

$$\mu_{11} = \frac{176700}{100} - (45 \cdot 8)(36 \cdot 6)$$

$$= 1767 - 1676 \cdot 28 = 90 \cdot 72$$

$$rxy = \frac{\mu_{11}}{\sigma x \cdot \sigma y} = \frac{90 \cdot 72}{(9 \cdot 559)(11 \cdot 93)}$$

$$= \frac{90 \cdot 72}{114 \cdot 039} = 0 \cdot 7955$$

x-மாறியை y-மாறியில் குறிக்கும் வளைகோடு

$$(x_i - \overline{x}) = \frac{\mu_{11}}{\sigma x^2} (y_i - \overline{y})$$

$$(x_i - 45 \cdot 8) = \frac{90 \cdot 72}{114 \cdot 039} (y^i - 36 \cdot 6)$$

$$x_i = 0 \cdot 6369 \; y_i + 45 \cdot 8 - 23 \cdot 31$$

$$x_i = 0 \cdot 6369 \, y_i + 22 \cdot 49$$

y-மாறியை x-மாறியில் குறிக்கும் வளைகோடு

$$(y_i - \overline{y}) = \frac{\mu_{11}}{\sigma y^2} (x_i - x)$$

$$(y_i - 36 \cdot 6) = \frac{90 \cdot 72}{91 \cdot 36} (x - 45 \cdot 8)$$

$$y_i = 0 \cdot 993 x_i + 36 \cdot 6 - 45 \cdot 48)$$

$$y_i = 0 \cdot 993 x_i - 8 \cdot 88$$

சுலபமுறை :

x, y மாறிகளை, சிறந்த தோராய மதிப்புகள், தகுந்த அலகுகளைக் கொண்டு u, v என்ற மாறிகளாக மாற்றித் தொடர்பு பட்டியலில் ஏற்படும் கணக்கியல் முறைகளைச் சுலபமாக்கிக் கொள்ளலாம். u, v என்ற மாறிகளாக மாற்றியபின் கணக்கிடப்படும் தொடர்பு கெழு கொடுக்கப்பட்ட மாறியின் தொடர்பு கெழுவிற்குச் சமமாக அமைவதைக் காணலாம்.

$$u = \frac{x_i - A}{c}$$

$$v = \frac{y_i - B}{d}$$

A, B என்பன x, y மாறிகளிலிருந்து கழிக்கப்படும் தோராய மதிப்புகளையும், c, d என்பன இவ் விலக்கங்கள் வகுக்கப்படும் அலகுகளையும் குறிக்கும். A, B என்ற தோராய மதிப்புகள் நிகழ்வெண் பரவலின் x, y மாற்றிகளின் பிரிவெண்களில் ஏதாவதொரு மதிப்பாகவும், c, d என்பன x, y மாறிகளின் இடைவெளி அளவுகளையும் குறிக்கும்.

$$x_i = A + c\,u_i$$

$$\bar{x} = A + c\,\bar{u}$$

$$y_i = B + d v_i$$

$$\bar{y} = B + d\bar{v}$$

$$rxy = \frac{\frac{1}{n}\Sigma\,(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n}\Sigma\,(x_i - \bar{x})^2 \cdot \frac{1}{n}\,\Sigma\,(y_i - \bar{y})^2}}$$

$$= \frac{\frac{1}{n}\,\Sigma\,(A + cu_i - A - c\bar{u})\;(B + dv_i - B - d\bar{v})}{\sqrt{\frac{1}{n}\,\Sigma\,A + c\bar{u}_i\;(-A - cu\,)^2\;\frac{1}{n}\,\Sigma\,(B + dv_i - B - d\bar{v}\,)^2}}$$

$$= \frac{\frac{1}{n}\Sigma\,c\,(u_i - \bar{u}\,)\;.\;d\;(v_i - \bar{v})}{\sqrt{\frac{1}{n}\;\Sigma c^2\,(u_i - \bar{u}\,)^2\;\frac{1}{n}\,\Sigma\,d^2\,(v_i - \bar{v}\,)^2}}$$

$$= \frac{\frac{1}{n} c \cdot d \Sigma (u - \overline{u})^2 \quad (v_i - \overline{v})}{c \cdot d\sqrt{\frac{1}{n} \Sigma \ (u_i - \overline{u}\ )^2 \ \frac{1}{n} \ \Sigma \ (v_i - \overline{v}\ )^2}}$$

$$= \frac{\frac{1}{n} \Sigma \ (u_i - \overline{u}\ ) \ (v - \overline{v}\ )}{\sqrt{\frac{1}{n} \ \Sigma \ (u - u) \frac{1}{n} \Sigma \ \ (v - v \ )^2}}$$

$$= ruv.$$

$$\therefore \ rxy = ruv.$$

இரு மாறிகளுக்கான உடன்தொடர்பைக் கணக்கிடும்போது ஏற்படும் சிரமத்தைக் குறைக்க தக்கத் தோராய மதிப்புகளைக் கழித்தும், சிறந்த அலகுகளால் வகுத்தும் அமைக்கப்படும் மாறிகளுக்கான தொடர்பு கெழு கொடுக்கப்பட்ட மாறிகளுக்கான தொடர்பு கெழுவிற்குச் சமமாகிறது. கூட்டுச் சராசரித் திட்டவிலக்கம் கணக்கிடும்போது மாற்றியமைக்கப்படும் மாறிகளின் மதிப்புகள் கணக்கிட்ட அளவைக்கேற்றவாறு கொடுக்கப்பட்ட மாறியில் மாற்றியமைக்கப்படவேண்டும் என்பதை அறிவோம். உடன்தொடர்பு கெழுவில் இத்தகைய மாற்றம் தேவையில்லை என்பதையே இது காட்டுகிறது.

உடன்தொடர்பு கெழுவைக் கணக்கிட உதவும் மதிப்புகளால் மாறிகளின் தொடர்பு கோடுகளும் கணக்கிடப்படும் என்பதை அறிவோம். மாற்றப்பட்ட மாறிகளின் மூலமே தொடர்பு கோடுகளும் குறிக்கப்படுவது தவறில்லையாயினும், மாறிகளின் தொடர்பு கோடுகளைக் கொடுக்கப்பட்ட மாறியில் குறிப்பிடுவது இந் நேர்கோட்டுச் சமன்பாடுகளை உபயோகப்படுத்தி மற்ற மாறியின் மதிப்பை நேரிடையாகக் கணக்கிட மிகவும் உபயோகமாக உள்ளது. கூட்டுச் சராசரி, திட்டவிலக்க அளவைகள் கொடுக்கப்பட்ட மாறிகளுக்கு மாற்றப்பட்டு மாறிகளின் தொடர்பு கோடுகள் கணக்கிடப்படுகின்றன. அடியிற்காணும் தொடர்பு பட்டியலில் இம்முறை பயன்படுத்தப்பட்டிருப்பதைக் காணலாம்.

உடன் தொடர்புப் பட்டியல்

$$\overline{u} = \frac{-1876}{1000} = -1{\cdot}876$$

$$\overline{v} = \frac{-1371}{1000} = -1{\cdot}371$$

$$\sigma u = \sqrt{\frac{5296}{1000} - (-1{\cdot}876)^2}$$

$$= \sqrt{5.296 - 3.5194} = \sqrt{1.7766}$$
$$= 1.332$$

$$\sigma v = \sqrt{\frac{3273}{1000} - (-1.371)^2}$$
$$= \sqrt{3.273 - 1.8796} = \sqrt{1.3934}$$
$$= 1.18$$

$$\mu_{11} = \frac{3727}{1000} - \left(\frac{-1876}{1000}\right)\left(\frac{-1371}{1000}\right)$$
$$= 3.727 - 2.572$$
$$= 1.155$$

$$r_{uv} = \frac{\mu_{11}}{\sigma u \cdot \sigma v} = \frac{1.155}{(1.332)(1.155)}$$
$$= 0.7347$$

மாறியின் தொடர்பு கோடுகள்

$$\bar{x} = A + c\bar{u}$$
$$= 37.5 + (5)(-1.876)$$
$$= 37.5 - 9.88 = 28.12$$
$$\bar{y} = B + d\bar{v}$$
$$= 32.5 + (5)(-1.371)$$
$$= 32.5 - 6.85 = 25.65$$
$$\sigma x = c \cdot \sigma u = 5 \times 1.332$$
$$= 6.66$$
$$\sigma y = c \cdot \sigma v = 5 \times 1.18$$
$$5.90$$

x-மாறியை y-மாறியில் குறிப்பிடும் நேர்கோடு

$$(x_i - \bar{x}) = r \frac{\sigma x}{\sigma y} (y_i - \bar{y})$$

$$(x_i - 28.12) = 0.7347 \cdot \frac{6.66}{5.90} (y - 25.65)$$
$$x_i = 0.8293 y_i + 28.12 - 21.27$$
$$x_i = 0.8293 y_i + 6.85$$

y-மாறியை x-மாறியில் குறிப்பிடும் நேர்கோடு

$$(y_i - \bar{y}) = r \cdot \frac{\sigma y}{\sigma x} (x_i - \bar{x})$$

$$(y_i - 25 \cdot 65) = 0 \cdot 7347 \frac{5 \cdot 90}{6 \cdot 66} (x - 28 \cdot 12)$$

$$y_i = 0 \cdot 651\, x - 18 \cdot 31 + 25 \cdot 65$$

$$y_i = 0 \cdot 651\, x_i + 7\ 34$$

உடன்தொடர்பு பட்டியல் (212ஆம் பக்கம் பார்க்க)

காட்டு :

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{2170}{1112} = 1\ 95$$

$$\bar{y} = \frac{\Sigma y_i}{n} = \frac{1881}{1112} = 1 \cdot 692$$

$$\sigma x = \sqrt{\frac{1}{n} \Sigma x_i^2\ f_i - \bar{x}^2}$$

$$\sigma x^2 = \frac{5350}{1112} - (1 \cdot 95)^2$$

$$= 4 \cdot 8112 - 3 \cdot 8025$$

$$= 1 \cdot 0087$$

$$\sigma y = \sqrt{\frac{1}{n} \Sigma 1_i^2\ f_i - \bar{y}^2}$$

$$\sigma y^2 = \frac{4295}{1112} - (1 \cdot 692)^2$$

$$= 2 \cdot 8948 - 3 \cdot 2994$$

$$= 0 \cdot 9995$$

உடன்தொடர்பு பட்டியல் (207ஆம் பக்கத்திற்குரியது)

y \ x	15–20	20–25	25–30	30–35	35–40	40–45	45–50	50–55	55–60		y_i	v_i	$v_i f_i$	$v_i^2 f$	$v_i u_i$
15–20	11 132	62 558	19 114	3 9	1 0					96	17·5	–3	–288	864	813
20–25	6 248	220 1320	190 760	34 68	6 0	2 –4				458	22·5	–	–916	1832	2192
25–30		46 138	165 330	59 59	13 0	5 –5	2 –4	1 –3		291	27·5	–1	–291	291	515
30–35		3 0	25 0	33 0	14 0	6 0	3 0	2 0		86	32·5	0	0	0	0
-40		1 –3	3 –6	8 –8	9 0	6 6	4 8	3 9	1 4	35	37·5	1	35	35	10
40–45				1 2	3 0	5 10	4 16	3 18	3 16	18	42·5	2	36	72	58
45–50				.	1 0	2 6	3 18	3 27	2 24	11	47·5		33	99	75

50–55							1	2 24	2 32	5	52·5	4	20	80	64
	17	332	402	138	47	26	17	14	7	1000			—1371	3273	3727
x_i	17·5	22·5	27·5	32·5	37·5	42·5	47·5	52·5	57·6						
u_i	—4	—3	—2	—1	0	1	2	3	4						
$u_i f_i$	—68	—996	—804	—138	0	26	34	42	28	—1816					
$u_i^2 f_i$	272	2988	1608	138	0	26	68	84	112	5296					
$u_i v_i$	180	2013	1198	126	0	13	46	75	76	3727					

உடன் தொடர்பு பட்டியல் (209 ஆம் பக்கத்திற்குரியது)

F / M	0	1	2	3	4	5	6	மொத்தம்	$y_i f_i$	$y_i^2 f_i$	$x_i y_i$
0			1 0	54 0	15 0	3 0	2 0	75	0	0	0
1		2 2	366 732	45 135	16 64	9 45	2 12	440	440	440	990
2	1 0	277 544	91 364	40 240	14 112	4 40	3 36	430	860	1720	1346
3	20 0	41 123	39 234	10 90		6 90		116	348	1044	537
4	5 0	9 36	10 80	2 24	2 32	1 20		29	116	464	192
5	3 0	6 30	4 40	2 30				15	75	375	100
6	1 0	3 18	3 36					7	42	252	54
மொத்தம்	30	338	514	153	47	23	7	1112	1881	4295	3219
$x_i f_i$	0	338	1028	459	188	115	42	2170			
$x_i^2 f_i$	0	338	2056	1377	752	575	252	5350			
$x_i f_i$	0	763	1486	519	208	195	48	3219			

$$\mu_{11} = \frac{1}{n} \Sigma\, x_i\, y_i - \bar{x}\, \bar{y}$$

$$= \frac{3219}{1112} - (1 \cdot 95)\,(1 \cdot 692)$$

$$= 2 \cdot 8948 - 3 \cdot 2994$$

$$= -0 \cdot 4046$$

$$r_{xy} = \frac{\mu_{11}}{\sigma x \cdot \sigma y} = \frac{-0 \cdot 4046}{1 \cdot 004}$$

$$= -\,0 \cdot 403$$

x-மாறியை y-மாறியில் குறிக்கும் நேர்கோடு

$$(x_i - \bar{x}) = \frac{\mu_{11}}{\sigma y^2}\,(y_i - \bar{y})$$

$$(x_i - 1 \cdot 95) = \frac{-0 \cdot 4046}{0 \cdot 9996}\,(y_i - 1 \cdot 692)$$

$$x_i = -\,0 \cdot 4081\, y_i + \cdot 6905 + 1 \cdot 95$$

$$= 2 \cdot 6405 - 0 \cdot 4081\, y_i$$

y-மாறியை x-மாறியில் குறிக்கும் நேர்கோடு

$$(y_i - 1 \cdot 692) = \frac{-0 \cdot 4046}{1 \cdot 0087}\,(x_i - 1 \cdot 95)$$

$$y_i = -\,0 \cdot 401\, x_i + 1 \cdot 692 + 0 \cdot 782$$

$$y_i = 2 \cdot 474 - 0 \cdot 401\, x_i$$

மதிப்பிடத் தொடர்பு கெழு (Rank correlation)

x, y-மாறிகளின் அளவுகளுக்கு மாறாக அவற்றின் தரம் அல்லது மதிப்பிடம் (rank) குறிப்பிடும்போது கணக்கிடப்படும் தொடர்புகெழு மதிப்பிடத் தொடர்பு கெழு எனப்படும். அளவைகளைவிட அவற்றின் மதிப்பிடம் குறிப்பிடுவதே சிறந்த தெனும் நிலையிலும், திட்டவட்டமாக அளவிட இயலாதனவான பண்புகளின் தன்மை மதிப்பிடங்களால் கொடுக்கப்படும் இடங் களிலும் இக் கெழு கணக்கிடப்படுகிறது. இரு மாறிகளுக்கான மதிப்புகளை $x_1\, x_2 \ldots\ldots x_n$, $y_1\, y_2 \ldots\ldots y_n$ எனக் கொள்வோம். இவை தரங்களாகவேனும் (மதிப்பிடங்கள்) குறிக்கப்படலாம். அளவு களாக உள்ள இடத்து அவற்றின் மதிப்புகளுக்குத் தக்கவாறு தரப்படுத்திக்கொள்ள வேண்டும்.

x_i	y_i	மதிப் பிடம் x_i	மதிப் பிடம் y_i	மதிப்பிட வேறுபாடு	மதிப்பிட வேறுபாட்டு வர்க்கம்
x_1	y_1	R_{x_1}	R_{y1}	$R_{x_1} - R_{y_1} = d_1$	d_1^2
x_2	y_2	R_{x_2}	R_{y_2}	$R_{x_2} - R_{y_2} = d_2$	d_2^2
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
x_n	y_n	R_{xn}	R_{yn}	$R_{xn} - R_{yn} = d_n$	d_n^2
					Σd_1^2

இரு மாறிகளுக்கான தரங்களின் வித்தியாச வர்க்கங்களின் கூட்டுத்தொகை கணக்கிடப்பட்டு (Σd_1^2) மதிப்பிடத் தொடர்பு கெழு வரையறுக்கப்படுகிறது.

$\rho = 1 - \dfrac{6 \Sigma d_i^2}{n(n^2 - 1)}$ என்ற வாய்பாட்டால் மதிப்பிடத் தொடர்பு கெழு குறிப்பிடப்படுகிறது. ρ (ரோ) என்பது மதிப்பிடத் தொடர்புகெழுவினையும், n என்பது உறுப்புகளின் எண்ணிக்கையையும் குறிக்கும். ஒவ்வொரு சோடி ($x_1\ y_1$, x_2, $y_2 \cdots x_n\ y_n$) உறுப்புகளின் மதிப்பிடங்கள் சமமாக இருக்கும்போது மதிப்பிடத் தொடர்பு கெழுவிற்கான மதிப்பிட வேறுபாட்டு வர்க்கம் பூச்சியமாகி $\rho = 1$ என அமையும். மதிப்பிட மதிப்புகள் நேர்மாறாக இருக்கும்போது $\rho = -1$ என அமையும். இவ்விரு எல்லைகளுக்கிடையில் அமையும் மதிப்பிடத் தொடர்பு கெழுவின் மதிப்பிற்குத் தக்கவாறு இதன் மதிப்பு விளக்கப்படும்.

காட்டு :

அழகு போட்டியில் இரு நடுவர்களால் குறிப்பிடப்பட்ட மதிப்பிடங்கள் தரப்பட்டுள்ளன. அழகினை மதிப்பிடும் தன்மையில் இரு நடுவர்களுக்குள்ள தொடர்பினைக் கணக்கிடு.

முதல் நடுவர் தந்த மதிப்பிடம்	இரண்டாம் நடுவர் தந்த மதிப்பிடம்	மதிப்பிட வேறுபாடு	மதிப்பிட வேறுபாட்டு வர்க்கம்
3	6	−3	9
5	4	+1	1
8	9	−1	1
4	8	−4	16
7	1	+6	36
10	2	+8	64
	3	−1	1
1	10	−9	81
6	5	+1	1
9	7	+2	4
$n = 10$			214

$$n = 10 \qquad \Sigma d_i^2 = 214$$

$$\rho = 1 - \frac{6 \Sigma d_i^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6 \times 214}{10 \cdot 99}$$

$$= 1 - 1{\cdot}297$$

$$\rho = -0{\cdot}297$$

இரு நடுவர்களும் சமமான மதிப்பிடங்கள் கொடுத்தவர்கள் என்று சொல்ல முடியாது. இருவரது அபூசின் மதிப்பிடும் தன்மையில் வேறுபாடு இருப்பதாகக் கொள்கிறோம்.

காட்டு :

அடியிற்காணும் அளவுகளுக்கான மதிப்பிடத் தொடர்பு கெழுவைக் கணக்கிடு.

$$x = \quad 57 \quad 33 \quad 45 \quad 50 \quad 48$$
$$y = \quad 20 \quad 60 \quad 58 \quad 54 \quad 56$$

x	y	மதிப்பிடம் x	மதிப்பிடம் y	மதிப்பிட வேறுபாடு	மதிப்பிட வேறு பாட்டு வர்க்கம்
57	20	1		−4	16
33	60	5	1	+4	16
45	58	2	4	+2	4
50	54	4	2	−2	4
48	56	3	3	0	0
					40

$$n = 5$$
$$\Sigma d_i^2 = 40$$
$$\rho = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)}$$
$$= 1 - \frac{6 \times 40}{5(25 - 1)}$$
$$= 1 - \frac{240}{120}$$
$$= 1 - 2$$
$$\rho = -1{\cdot}0$$

ஒரு மாறியின் இருஅளவுகள் ஒரேமதிப்பைப் பெறும்போது அவற்றிற்கான மதிப்பிடம் தருவதில் எழும் சிக்கலைப் போக்க அடியிற்காணும் முறை கையாளப்படுகிறது. 4 மதிப்பிடங்கள் வரையறுக்கப்பட்டபின், இருஅளவுகள் சமமாக இருப்பின்

அவற்றின் மதிப்பிடம் 5,6 எனக்கொண்டு இரு அளவுகளுக்குமான மதிப்பிடங்கள் 5·5 எனக் குறிக்கப்படும். இதை அடுத்துவரும் மூன்று அளவுகள் சமமாக இருப்பின், இவற்றிற்கான மதிப்பிடங்கள் 7,8,9 எனக் கொண்டு, இவற்றின் கூட்டுச் சராசரியான 8 ஆனது $\left(\frac{7+8+19}{3}\right)$ ஒவ்வொரு மதிப்பிற்குமான மதிப்பிடங்களாகக் குறிக்கப்படும். எவ்வளவு அளவுகள் சமமாக அமைகிறதோ, அவ் வளவுகளின் மதிப்பிடங்களின் சராசரி ஒவ்வோர் இடத்திற்குமான மதிப்பிடமாகக் குறிக்கப்படும். எல்லா அளவுகளும் மாறுபட்ட மதிப்பிடங்களைப் பெறும் என்ற அடிப்படையில் மதிப்பிடத் தொடர்பு கெழுவின் வாய்பாடு வரையறுக்கப்படுவதால், இரண்டு அல்லது அதற்கு மேற்பட்ட அளவுகள் சமமான மதிப்பைப் பெறும்போது நம்மால் கையாளப்படும் மேற்கண்ட முறையில் ஏற்படும் பிழைகளைச் சரிசெய்ய Σd_i^2 உடன் $\frac{1}{12}(m^3-m)$ என்ற மதிப்புக் கூட்டப்படுகிறது. m என்பது சமமான மதிப்புகள் பெற்ற அளவுகளின் எண்ணிக்கையைக் குறிக்கும். x-மாறியில் இரண்டளவுகளும், y-மாறியில் மூன்றளவுகளும் சமமான மதிப்பைப் பெற்றிருப்பின்,

$$\rho = 1 - \frac{6\,\Sigma d_i^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m)}{n(n^2 - 1)}$$

என்ற மதிப்பு

$$\rho = 1 - \frac{6\,\Sigma d_i^2 + \frac{1}{12}(2^3 - 2) + \frac{1}{12}(3^3 - 3)}{n(n^2 - 1)}$$

என அமையும்.

காட்டு :

அடியிற்காணும் அளவுகளுக்கான மதிப்பிடத் தொடர்பு கெழுவைக் கணக்கிடுவோம்.

$x =$	38	23	30	4	6	6	55	14	6	47
$y =$	8	8	19	1	10	10	15	4	1	14

x	y	மதிப்பிடம் x	மதிப்பிடம் y	மதிப்பீட வேறுபாடு	மதிப்பீட வேறுபாட்டு வர்க்கம்
38	8	5·5	3	—2·5	6·25
23	8	5·5	5	—0·5	0·25
30	19	1	4	3·0	9·00
4	1	8·5	10	1·5	2·25
6	10	4	8	4·0	16·00
6	10	10	8	—2·0	4·00
55	15	2	1	—1·0	1·00
14	4	7	6	—1·0	1·00
6	1	8·5	8	—0·5	0·25
47	14	3	2	—1·0	1·00
					41·00

$$\rho = 1 - \frac{6\,\Sigma\, d_i^2 + \frac{1}{12}(m^3-m) + \frac{1}{12}(m^3-m) + \frac{1}{12}(m^3-m)}{n(n^2-1)}$$

$$= 1 - \frac{6\,[41 + \frac{1}{12}(3^3-3) + \frac{1}{12}(2^3-2) + \frac{1}{12}(2^3-2)}{10\,(10^2-1)}$$

$$\rho = 1 - \frac{6\;[41 + 2 + 0\cdot5 + 0\cdot5]}{990}$$

$$= 1 - \frac{264}{990}$$

$$= 1 - 0\cdot2667$$

$$\rho = 0\cdot7333$$

பயிற்சி :

1. அடியிற்காணும் அளவுகளுக்கான உடன்தொடர்பு கெழுவைக் கணக்கிடு.

x	1	3	4	2	5	8	9	10	13	15
y	8	6	10	8	12	16	16	10	32	32

2. கம்பு, சோளம் ஆகியவற்றின் விலைகளுக்கான சிதறல் விளக்கப்படம் வரைக.

வருடம்	விலை 10 மூட்டைகளுக்கானது (ரூபாயில்)		வருடம்	விலை 10 மூட்டைகளுக்கானது (ரூபாயில்)	
	கம்பு	சோளம்		கம்பு	சோளம்
1905–06	279	298	1910–11	281	318
1906–07	295	315	1911–12	317	353
1907–08	326	338	1912–13	337	396
1908–09	346	369	1913–14	337	383
1909-10	315	334	1914–15	309	326

3. பிற்காணும் பட்டியலில், மொத்த மக்கள்தொகையும், அவர்களின் பகுப்பில் அமையும் குறைபார்வை உடையோர்தம் எண்ணிக்கையும் கொடுக்கப்பட்டுள்ளன. வயதிற்கும், குறைபார்வைக்கும் உள்ள தொடர்பைக் கணக்கிடு.

வயது	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
மக்கள் தொகை (ஆயிரத்தில்)	100	60	40	36	24	11	6	3
குறை பார்வையுடையோர்	45	40	40	40	36	22	18	15

4. உயிரியலிலும், உயிர்ப் புள்ளியியலிலும் 11 மாணவர்கள் தம் மதிப்பெண்கள் தரப்பட்டுள்ளன. இரு மாறிகளுக்குமான உடன்தொடர்பைக் கணக்கிடு.

உயிரியல் (100-க்கு)	45	55	56	58	60	65	68	70	75	80	85
உயிர்ப் புள்ளியியல் 100-க்கு	56	50	48	60	62	64	65	70	74	82	90

5. நீரின் அளவிற்கும், அரிசி விளைவிற்குமான அளவுகள் அடியிற்காணும் பட்டியலில் தரப்பட்டுள்ளன. நீரின் அளவிற்கான அரிசியின் விளைபலன் தொடர்புகோட்டினை அமை. 40 அங்குல நீர் கிடைக்கும்போது ஏற்படக்கூடிய விளைவின் அளவைக் கணக்கிடு.

நீரின் அளவு (அங்.)	12	18	24	30	36	42	48
விளைபலன் (டன்னில்)	5·27	5·68	6·25	7·21	8·02	8·71	8·42

6. 90 இடங்களின் மழை அளவிற்கும், வெப்பநிலைக்குமான தொடர்பு பட்டியல் கொடுக்கப்பட்டுள்ளது. இவ்விருமாறிகளுக்குமான தொடர்புகெழு, தொடர்பு கோடுகளைக் கணக்கிடு.

வெப்பநிலை

மழை அளவு

x	10–14	15–19	20–24	25–29	30–34	35–39	மொத்தம்
16-18					1	1	2
19-21	1		2	2	1	2	8
22-24			1	2	1		4
25-27	2	5	6	6	3	1	23
28-30	3	8	4	4	2		21
31-33	4	6	4	3			17
34-36	2	4	3	2			11
37-39	1	2	1				4
மொத்தம்	13	25	21	19	8	4	90

7. கொடுக்கப்பட்டுள்ள அளவுகளுக்கான தொடர்புகெழு மாறியின் தொடர்பு கோடுகளைக் கணக்கிடு.

x	1	2	3	4	5	6	7	8	9
y	9	8	10	12	11	13	14	16	15

8. தந்தை, மகன் உயரங்களின் அளவுகளுக்கான தொடர்பு கெழுவையும், மாறிகளின் தொடர்பு கோடுகளையும் கணக்கிடு.

தந்தை உயரம்	65	66	67	67	68	69	71	72	73
மகன் உயரம்	67	68	64	68	72	70	70	69	70

9. நெல்விளைவைக் காணச் செய்த சோதனை ஒன்றில், அறுக்கப்பட்ட உடனே எடுக்கப்பட்ட எடைக்கும், காய்ந்தபின் எடுக்கப்பட்ட எடைக்குமான ‘இருமாறி நிகழ்வெண் பட்டியல்’ (Bivariate frequency table) பின்வருமாறு

காய்ந்தபின் எடுக்கப்பட்ட எடை

அறுத்தவுடன் எடுக்கப்பட்ட எடை

x \ y	20—25	25—30	30—35	35—40	40—45	45—50
35–40	1	2	2			
40–45	1	2	3	2		
45–50		3	5	3	3	1
50–55			4	6	5	1
55–60			1	1	4	

10. ஜன நெருக்கடிக்கும் குழந்தை இறப்பு விகிதத்திற்குமான அளவுகளிலிருந்து அவற்றுக்கிடையில் உள்ள தொடர்பைக் கணக்கிடு.

குழந்தை இறப்பு விகிதம்	109	122	96	142	151	124	125	102	109	156	122
ஜன நெருக்கடி	14·9	6·3	5·8	12·2	33·2	13·3	14·6	8·8	4·9	39·8	6·3

11. பத்துச் சோடி புதுமணத் தம்பதிகளுக்கான வயதில் உள்ள உடன்தொடர்பைக் கணக்கிடு.

கணவன் வயது	17	18	19	19	20	20	21	21	22	23
மனைவி வயது	12	16	14	11	15	19	22	16	15	20

12. 1000 நபர்களின் எடை, உயரத்திற்கான தொடர்பு பட்டியல் கொடுக்கப்பட்டுள்ளது. எடைக்கும் உயரத்திற்கும் உள்ள மாறிகளின் தொடர்பு கோடுகளைக் கணக்கிடு. 149 பவுண்டு எடையுள்ளவரின் சராசரி உயரத்தையும், 66 அங். உயரம் உள்ளவர்களின் சராசரி எடையையும் கணக்கிடு.

x எடை / y உயரம்	85—100	100—115	115—130	130—145	145—160	160—175	175—190	190—205	மொத்தம்
60-62	3	9	1						13
62-64	9	32	40	20	8	1			110
64-66	7	39	95	51	31	4			227
66-68	2	25	104	110	60	8	2		311
68-70	1	4	42	71	49	40	5	2	214
70-72		1	5	11	32	35	4		88
72-74			1	1	5	19	3	2	31
74-76				1		2	2	1	6
மொத்தம	22	110	288	265	185	109	16	5	1000

13. 1876 முதல் 1930 வரை, ஒரு நாட்டின் பிறப்பு இறப்பு விகிதங்கள் கொடுக்கப்பட்டுள்ளன. பிறப்பு இறப்பு விகிதங்களுக்கான உடன்தொடர்பைக் கணக்கிடு.

வருடம்	பிறப்பு விகிதம்	இறப்பு விகிதம்
1876—80	35 3	20·8
1881—85	33 5	19·4
1886—90	31·4	18·9
1891—95	30 5	18·7
1896—1900	29·3	17·7
1901—1905	28·2	16·0
1906—1910	26·3	14·7
1911—1915	23 6	14·3
1916—20	20·6	14·4
1921—25	19 9	12·2
1926—1930	16·7	12·1

14. அடியிற்காணும் அளவுகளுக்கான மாறிகளின் தொடர்பு கோடுகளைக் கணக்கிடு.

x	152	114	138	154	144	153	141	117	136	154
y	193	300	414	594	676	549	320	483	481	659

15. பின்வரும் அளவைகளுக்கான மாறிகளின் தொடர்பு கோடுகளை மதிப்பிடு.

$\bar{x} = 125 \qquad \sigma x = 15$

$\bar{y} = 80 \qquad \sigma y = 9 \qquad r_{xy} = 0.55$

16. 10 மாணவர்கள் இரு பாடங்களில் பெற்ற மதிப்பெண்கள் கொடுக்கப்பட்டுள்ளன. இவற்றிற்கான மதிப்பிடத் தொடர்புகெழுவைக் (rank correlation) கணக்கிடு.

x	75	45	60	16	25	25	90	35	28	85
y	4	40	90	10	50	0	75	25	12	70

17. குறை உணவு கொடுக்கப்பட்ட ஏழு எலிகளின் உடல் நிலையை இரு ஆராய்ச்சியாளர்கள் ஒருவாரத்திற்குப் பின்பு சோதித்துக் குறித்த மதிப்பிடங்கள் (rank) தரப்பட்டுள்ளன. இதற்கான மதிப்பிடத் தொடர்பு கெழுவினைக் கணக்கிடு.

எலி எண்	மதிப்பிடம்	
	ஆராய்ச்சியாளர் 1	ஆராய்ச்சியாளர் 2
1		4
2	1	2
3	6	5
4	5	6
5	3	1
6	2	3
7	7	7

18. ஒரு கல்லூரியில் உள்ள மாணவர்களின் உயரத்திற்கும் அவர்களின் மார்பளவிற்கும் உள்ள தொடர்பு பட்டியல் கொடுக்கப்பட்டுள்ளது. உயரத்திற்கும் மார்பளவிற்கும் உள்ள தொடர்பினைக் கணக்கிடு.

உயரம் / மார்பளவு	60″	61″	62″	63″	64″	65″	66″	67″	68″	69″	70″	71″
24	1											
25		1										
26	2		1									
27		2	2	1	1							
28	3	1	1	5	1	1	2	1				
29		3	7	4	5	7	7	3				
30	1	4	7	8	15	9	8	7	4	3		
31		1	3	4	6	6	10	6	1			
32		4	5	5	7	4	7	4	3	1	1	1
33			1	3	4	2	4	3	1			
34	1	1					2	3				
35					2		2	1				
36					1	1						
37					1	1						

19. அடியிற்காணும் அளவுகளுக்கான தொடர்பு பட்டியலை அமைத்து, தொடர்பு கெழு, மாறிகளின் தொடர்பு கோடுகளைக் கணக்கிடு.

$x =$	32	17	10	8	9	11	33	20	34	31
$y =$	12	16	30	40	36	26	14	13	12	13
$x =$	15	18	15	21	16	12	14	19	27	29
$y =$	18	15	20	14	17	22	18	13	12	12

20. 10 மாணவர்களின் எடை, உயரத்திற்கான மதிப்புகள் தரப்பட்டுள்ளன. அவற்றிற்கான மாறிகளின் தொடர்பு கோடுகளை அமை. 71·5 அங். உயரமுள்ள மாணவனின் எடையையும், 146 பவுண்டு எடையுள்ள மாணவனின் உயரத்தையும் கணக்கிடு.

x	63	72	70	68	66	69	74	70	63	72
y	124	184	161	164	140	154	210	164	126	172

9. கூறெடுத்தல்
(Sampling)

மாதிரி, மாதிரியின் பரவல், திட்டப்பிழை
(Sample, Sampling distribution, Standard error)

(பொதுவான கருத்துகள்)

பெரும்பாலான இடங்களில் முழுமைத்தொகுதியிலுள்ள (Population) எல்லா உறுப்புகளின் பண்பினைத் தெரியாமலேயே, முழுமைத் தொகுதிக்கான அப் பண்பினைப்பற்றிய முடிவுகளை எடுக்கவேண்டியவர்களாக உள்ளோம். நாள்தோறும் நம்மை அறியாமலேயே நடைமுறையில் இப் பழக்கத்தைக் கடைப்பிடித்து வருகிறோம். நமக்குத் தேவையான துணி வாங்கும்போது அதன் ஒரு பகுதியை மட்டும் சோதனை செய்து, துணியின் மொத்தப் பண்பினைப் பற்றிய ஒரு முடிவிற்கு வருகிறோம். ஒரு மூட்டையிலிருந்து எடுக்கப்பட்ட ஒரு கையளவு அரிசி அம் மூட்டையிலுள்ள அரிசியின் தரத்தை மதிப்பிடும் என உணர்கிறோம். இன்னும் சில இடங்களில் முழுமைத் தொகுதி முழுவதையும் சோதனை செய்வது தேவையற்றது என எண்ணுகிறோம். ஒரு கிடங்கில் உள்ள உருளைக் கிழங்கின் தரத்தைக் காணக் கிடங்குக் கிழங்கு முழுதும் சோதனை செய்வது பொருளற்றதாகும். இன்னும் சில இடங்களில் முழுமைத் தொகுதி முழுவதையும் சோதனைக்குள்ளாவது முடியாத செயலாக அமைகிறது. பூச்சிகொல்லி மருந்து (Pesticide) ஒன்றின் தரம் காண மருந்து முழுவதும் சோதனைக்குள்ளாக்கப்பட்டால் விற்பனைக்கு மீதமாவது பூச்சியமாகும். ஒரு குறிப்பிட்ட அளவு மருந்தினைச் சோதனை செய்து, அதன் சக்தியைப்போலவே அம் முறைப்படி தயார் செய்யப்பட்ட மருந்து முழுவதும் இருக்குமென நம்புகிறோம். எனவே, முழுமைத் தொகுதியின் பண்புகளை அறிய அதிலிருந்து எடுக்கப்படும் ஒருபகுதியை ஆராய்ந்து முடிவிற்கு வருவது அறிவியலுக்கு ஒவ்வாது எனக் கொள்ள இடமில்லை.

முழுமைத் தொகுதியிலிருந்து ஒரு பிரதிநிதித்துவக் கூறினை, (representative) எடுத்து, அதற்கான பண்புகளை ஆராய்ந்து முழுமைத் தொகுதிக்கான பண்புகளைப் புள்ளியியல் முறைப்படி கூறுவதே கூறெடுத்தலின் நியதியாக உள்ளது. முழுமைத் தொகுதியையும் ஆராய்ந்து சந்தேகத்திற்கிடமான, வரையறுக்கப் படாத முடிவுகளைக் கூறுவதைக் காட்டிலும், அதன் ஒரு பகுதியினை (sample) ஆராய்ந்து நம்பகமான முடிவுக்கு வரமுடியும் என்பதே இதன் உட்கருத்தாகும். நிகழ்திறத்திற்கு அடுத்துப் புள்ளியியலின் அடிப்படைக் கருத்தாக அமைந்துள்ளது கூறெடுத்தல் பிரிவாகும். மதிப்பீடு (estimation) கருதுகோட் சோதனைகளுக்கான (Testing of Hypothesis) நம்பகத் தன்மை கூறெடுத்தலின் முறையினால் பாதிக்கப்படுவதால் இதன் முக்கியத் துவம் இன்னும் அதிகப்படுத்தப்படுகிறது. இதற்கான சில அடிப் படைக் கருத்துகளைக் காண்போம்.

முழுமைத் தொகுதி (population) என்பது நமது பிரச்சினைக்குரிய பொருள்கள் அல்லது அளவைகளின் ஏதாவது ஒரு பண்பு (attribute) அல்லது அளவு (measurement) ஆகியவற்றின் கூட்டினைக் குறிக்கும்; உயிர்ப்பொருள்கள் மட்டுமன்றிச் செயல்கள் மற்றும் பருப்பொருள்களையும் இது குறிக்கலாம். பொருள் களாலும், உள்பொருள்களாலும் (entity) அல்லது இத்தன்மைத் தான பொருள்களின் பண்புகளாலும் இதை வரையறுக்கிறோம். இடம், காலம் அல்லது நம்மால் குறிப்பிட முடியும் எந்த அடை யாளத்தாலும் இதன் தன்மை குறிக்கப்படலாம். முழுமைத் தொகுதியின் பொருள்களாக வயது வந்த ஆடவர்களையும், விதைப் பொட்டலங்களையும், விவசாயப் பண்ணைகளையும் அல்லது இம் முழுமைத் தொகுதியின் தனிப் பண்புகளான ஆடவர்தம் வயது, எடை, வருமானம், விதைகளின் எண்ணிக்கை, பழுதுபட்ட தன்மை, முளைக்கும் தன்மை, அவற்றில் உள்ள இனங்களின் விகித சமம், பண்ணையின் பரப்பு, பயிர் வகைகளின் இனம் ஆகிய வற்றைக் காட்டாகச் சொல்லலாம்.

முழுமைத் தொகுதிக்கான பொருள்கள் பல பண்புகளைக் கொண்டதாக இருப்பினும், நமது பிரச்சினை அதற்கான ஒன்று அல்லது இரண்டு பண்புகளைப்பற்றியதாகவே அமையும். குறைந்த பரப்பில் அமைந்ததும், மாறிலியானதும், இலகுவாக அடையாளம் காணும் தன்மைத்தானதுமான பொருள்களின் முழுமைத் தொகுதியின் எல்லைகளைக் காணுவது கடினமன்று. இவை குறிப் பிட்ட காலத்தில் வரையறுக்கப்பட்டு நிலையான தன்மைத்தாய் அமையும். இத்தகைய முழுமைத் தொகுதி நிலையான முழுமைத்

தொகுதி (static population) எனக் கொள்ளப்படும். முழுமைத் தொகுதியின் பொருள்கள் அடிக்கடி அழிக்கப்பட்டும், மாற்றப் பட்டும் புதுப் பொருள்கள் முழுமைத் தொகுதியில் சேர்வதாயும் இருப்பின், அது இயங்கும் தன்மைத்தான முழுமைத் தொகுதி (dynamic population) எனப்படும். ஒரு கடைக்கான வாடிக்கைக் காரர்களையும், ஓடையில் உள்ள நீரினையும் இயங்கும் தன்மைத் தான முழுமைத்தொகுதிக்கும், மண்டியில் உள்ள கிழங்கு, கல்லூரி மாணவர்களின் எண்ணிக்கை (ஓர் ஆண்டில்) இவற்றை நிலையான தன்மைத்தான முழுமைத் தொகுதிக்கும் காட்டாகச் சொல்ல லாம்.

இயல்பான அல்லது ஊகமில்லாத முழுமைத்தொகுதி (real population) உண்மையான பொருள்களைக் கொண்டு உருப் பெற்றதாய் இருக்கும். செயற்கையான முழுமைத் தொகுதி (artificial population) என்பது ஆராய்ச்சியாளர்தம் கருத்துகளை விளக்கவோ அல்லது புள்ளியியல் கருத்துகளை இலகுவாக உணர்த்தவோ ஊகத்தால் அமைக்கப்பட்ட முழுமைத் தொகுதி யாகும். கருதுகோள் முழுமைத் தொகுதி (hypothetical population) என்பது புள்ளியியலாரால் கற்பனை செய்யப்பட்டு உண்மையில் அமைவதற்கான வாய்ப்பைக் கொண்டது என நம்பும் முழுமைத் தொகுதியாகும். நாணயம் அல்லது பகடையை வீசுவதால் கிடைக்கும் முழுமைத் தொகுதியை இதற்குக் காட்டாகச் சொல்லலாம்.

வரையறுக்கப்பட்ட முழுமைத் தொகுதி (finite population) யிலுள்ள பொருள்கள் எண்ணத் தகுந்தவைகளாகவும், மொத்த எண்ணிக்கை வரையறுக்கப்பட்டதாயும் அமைந்திருக்கும். தஞ்சை மாவட்டத்திலுள்ள பண்ணைகளையும், கோவை விவசாயப் பண் ணையில் அமைந்த சோதனைப் பாத்திகளையும் இதற்குக் காட்டாகச் சொல்லலாம். வரையறுக்கப்படாத அல்லது முடிவிலா முழுமைத் தொகுதியிலுள்ள (infinite population) உறுப்புகள் கணக்கற்றன வாகவோ அல்லது கணக்கிடமுடியாதனவாகவோ அமைந்திருக் கும். விண்மீன்களின் கூட்டத்தை இதற்குக் காட்டாகச் சொல்ல லாம். கடல் நீரிலிருந்து எடுக்கக்கூடிய மாதிரிகள் ஒரு முடிவில் லாத நிலையாய் உள்ளது. நமது தேவையை உணர்ந்து முழுமைத் தொகுதியின் எல்லையை நாமே வரையறுத்துக் கொள்ள வேண்டி யவர்களாக உள்ளோம். நமக்குரிய பிரச்சினையின் தன்மையாலும், பொருள் வசதியாலும், சிக்கலை அணுகும் நோக்கத்தாலும் முழுமைத் தொகுயின் எல்லைகள் வரையறுக்கப்படுகின்றன. காற்று, திரவம், தூள் ஆகியவற்றின் முழுமைத் தொகுதிகள் வரை

யறுக்கப்பட்டதெனக் கொள்வது கடினமாதலின், இதன் முழுமைத் தொகுதிகள் முடிவற்ற முழுமைத் தொகுதிகள் எனக் கொள்ளப்படும். சில முடிவுள்ள முழுமைத் தொகுதிகளின் அமைப்பைக் கொண்டு அவற்றிலிருந்து மாதிரிகள் எடுக்கும் போது, அம் முழுமைத் தொகுதிகளை முடிவற்ற (infinite) முழுமைத் தொகுதிகள் எனக்கொள்வது நடைமுறையில் இருந்து வருகிறது.

அறிந்த முழுமைத் தொகுதி (known population) என்பது நமது ஆய்வுக்குட்பட்ட பண்பிற்கான மாறியின் அடிப்படைத் தன்மையையும், அதற்கான நிகழ்வெண் சார்பினையும் (frequency function) தெரிந்தவர்கள் என்பதைச் சுட்டும். இத் தன்மைத்தாய் இல்லாத முழுமைத் தொகுதிகள் அறியாத முழுமைத்தொகுதிகள் (unknown population) எனப்படும். நிகழ்வெண் சார்பு தெரியாத, நமது ஆய்விற்குரிய மாறியின் முழுமைத் தொகுதிபற்றிய முழுமை அளவைகளை (parameter) மாதிரி அளவை (statistic) களைக் கொண்டு அளவிடுவதே புள்ளியியல் பிரிவின் முக்கியமானவொரு பிரச்சினையாக அமைந்துள்ளது.

மாதிரிகள் அனைத்தும் மேற்கண்ட ஏதாவதொரு பண்பினைப் பெற்ற முழுமைத் தொகுதியிலிருந்து நிகழ்திற நியதிக்குக் கட்டுப்பட்டு எடுக்கப்பட்ட பகுதிகளாகும். முழுமைத் தொகுதிக்கும் ஒரு சிறந்த பிரதிநிதியாக இம் மாதிரி அமையும்போது, நம்மால் எடுக்கப்படும் முடிவுகளின் நம்பும் தன்மை அதிகரிக்கிறது. மாதிரியின் உறுப்புகள் பிரதிநிதித்துவ நிலையை அடைந்ததாயும், இத் தன்மை அமைவதற்கு நிகழ்திற நியதி பயன்படுத்தப்பட்டு அமைவது அவசியம். இத் தன்மை அமைவதற்கான சில அடிப்படைக் கருத்துகளைக் காண்போம்.

இயைபிலா மாதிரி (Simple Random Sampling)

முழுமைத் தொகுதியில் உள்ள எல்லா உறுப்புகளும் மாதிரியில் சேர்க்கப்படுவதற்கான நிகழ்திறன் சமமாக இருக்கும்படி அமையப்பெற்ற மாதிரி ராண்டம் மாதிரி (random sample) அல்லது இயைபிலா மாதிரி எனப்படும். மாதிரியில் உள்ள எந்தப் பொருளும், மாதிரியில் சேர்க்கப்பட்ட மற்றப் பொருள்களைக் காட்டிலும் அதிக வாய்ப்பினைப் பெற்றதாய் இல்லாத நிலையை இது குறிக்கும். முழுமைத் தொகுதியிலுள்ள சில பொருள்கள் மாதிரியில் சேர்க்கப்படுவதற்கான நிகழ்திறன் செயற்கையாக அதிகப்படுத்தப்பட்டிருப்பின் நமக்குக் கிடைத்துள்ள மாதிரி ஒரு சார்புடைய மாதிரி (biased sample) எனப்படும். சார்பற்ற (unbiased)

மாதிரிகளை எடுப்பதற்கான சில யுத்திகளைக் கூறுவதே புள்ளியியலின் நோக்கமாகும்.

சார்பற்ற மாதிரி எடுக்கப்பட வேண்டுமெனச் சொல்வது மிக எளிதாயினும் செயல்முறையில் நிறைந்த சிக்கல்களை ஏற்படுத்துகிறது. ராண்டம் (random) என்ற சொல்லே புரிந்து கொள்ள முடியாத ஒரு நியதியைப் பெற்றுள்ளதை உணர்கிறோம்.

எல்லா உறுப்புகளுக்கும் சம வாய்ப்புக் கொடுப்பதால் கிடைக்கும் மாதிரி ராண்டம் மாதிரியாக இருக்குமெனக் கொள்வது புள்ளியியலின் ஒரு யுத்தியாகும். 30 வயல்களிலிருந்து 5 வயல்களை ராண்டமாக எடுத்து விளைபலனைக் காண ஓர் அதிகாரி திட்டமிடுவதாகக் கொள்வோம். மாதிரியில் அமைந்த 5 வயல்களும் முழுமைத் தொகுதியிலுள்ள 30 வயல்களுக்கும் சமவாய்ப்புக் கொடுக்கப்பட்டு தேர்ந்தெடுக்கப்படுவது இயைபிலா மாதிரியாகும். அதிகாரி தம் அனுபவத்தால்தான் இதைச் செய்ய முடியும். இதன் முடிவுகள் அவரை எப்படிப் பாதிக்கிறது என்பதைப் பொறுத்தும் அவர்தம் மாதிரி அமையலாம். மிகவும் நன்கு விளைந்த வயல்களையோ அல்லது நன்கு விளையாத வயல்களையோ தேர்ந்தெடுக்கப்படும் சூழல் அமைவதை உணர முடிகிறது. தனிப்பட்டவர்களால் எடுக்கப்படும் மாதிரிகள் இயைபிலாத் தன்மையைப் பெறுவது அரிதாக உள்ளது. புள்ளியியலில் அமைந்துள்ள பல முறைகளில், ராண்டம் எண்களைக் (random number) கொண்டு இயைபிலா மாதிரிகளை அமைப்பது சிறந்ததாக அமைந்துள்ளது.

0, 1, 2, 3, 4. 5, 6, 7, 8, 9 ஆகிய எண்கள் ஒவ்வொன்றும் சமமான எண்ணிக்கையில் அமைவதாயும், அப்படி அமையாத இடத்து, அவை அமைந்துள்ள முறை புள்ளியியல் முறைப்படி குறிப்பிடத் தக்கபடி அமையாத அமைப்பைப் பெற்றிருப்பதாயும் வரிசைப்படுத்தப்பட்ட எண்களே ராண்டம் எண்களாகும். முழுமைத் தொகுதியிலுள்ள உறுப்புகளை வரிசைப் படுத்தி எண்ணிட்டபின் ராண்டம் எண்களை உபயோகப் படுத்திச் சார்பற்ற மாதிரிகளைத் தேர்ந்தெடுக்கலாம்.

30 வயல்கள் உள்ள ஒரு பண்ணையிலிருந்து 5 வயல்கள் கொண்ட மாதிரியைத் தேர்ந்தெடுப்பதற்கு அடியிற் காணும் ராண்டம் எண்களை உபயோகப்படுத்தும் முறையைக் காண்போம்.

5525	2106	3403	4926	7858	3755
0135	1737	5504	0171	9765	6478
7403	4174	6749	1115	1256	4163
7578	3127	9779	1225	3014	8330
8030	6396	5630	9512	4867	1879

வயல்களின் உரிமையாளர் அகர வரிசை அல்லது வயலின் பரப்பிற்கான வரிசைப்படி 30 வயல்களையும் பகுத்து எண்ணிடுவோம். அவை முறையே

00	05	10	15	20	25
01	06	11	16	21	26
02	07	12	17	22	27
03	08	13	18	23	28
04	09	14	19	24	29

என அமையும். 00 என்ற அமைப்பு ராண்டம் எண்களில் அமையப் பெறுவதால், 00 என முதல் பண்ணையைக் குறிப்பிடுவது அவசியமாகிறது. ராண்டம் எண்களின் முதற் பாகத்தில் அமைந்த இரண்டு எண்களை வரிசையாகப் பார்த்து 30-க்குக் குறைவாக அமைந்த 5 மாறுபட்ட எண்களை எடுக்கவேண்டும். இவை முறையே 04, 24, 17, 01, 11 என அமைகின்றன. 30-க்கு அதிகமான மதிப்புள்ள எண்களையும், திரும்பவும் அமையும் எண்களையும் (repeated numbers) தள்ளிவிடுகிறோம். ஒரே பண்ணையை இருமுறை தேர்ந்தெடுக்க முடியாதாகையால், திரும்பவும் அதே எண் வரும்போது தள்ளிவிடுவது அவசியமாகிறது. எடுக்கப்பட்ட எண்களுக்கு எதிராகக் குறிக்கப்பட்ட வயல்களே மாதிரியில் சேர்க்கப்படும் வயல்களாக அமையும்.

சீட்டுக் கட்டொன்றிலிருந்து எடுக்கப்படும் சீட்டு, தட்டொன்றிலிருந்து எடுக்கப்படும் பந்து ஆகியவை முழுமைத் தொகுதி உறுப்புகளாக அமையும்போது முழுமைத் தொகுதிகள் எண்ணிக்கையை மாறிலியாக (constant) இருக்கச் செய்யலாம். எடுக்கப்பட்ட சீட்டு அல்லது பந்தை மறுபடியும் சேர்த்தபின் இரண்டாவது சீட்டு அல்லது பந்து எடுக்கப்படுகிறது. எண்ணிக்கையை இவ்வாறு ஒவ்வொரு தடவையும் சமன் செய்யும்போது, வரையறுக்கப்பட்ட முழுமைத் தொகுதியிலிருந்து எடுக்கப்படும் பொருள் அனைத்துக்கும் சமவாய்ப்புக் கொடுக்கும் நிலை உறுதிப்படுத்தப்படுகிறது. இதில் எடுத்த பொருளையே மீண்டும் எடுக்கும் அபாயம் நிறைந்துள்ளதை உணர்கிறோம். எடுத்த பொருளையே மீண்டும் எடுக்கும் நிலவரின் அதைத் தள்ளி மாதிரியில் அமைந்த

உறுப்புகள் யாவும் மாறுபட்ட தன்மைத்தாய் இருக்கச் செய்கிறோம். இதே முறையைத்தான் ராண்டம் எண்களை உபயோகப்படுத்தும்போதும் கடைப்பிடிக்கிறோம். ஒரே எண் மறுபடியும் நமது தேர்வில் அமைவது எடுத்த பொருளையே மீண்டும் எடுப்பதற்குச் சமமாகும். செயற்களத்தில் இது சாதசமானதாகவோ அல்லது முடியாததாகவோ அமைந்துவிடுகிறது. குறிப்பாக முதலில் எடுக்கப்படும் பொருள், சோதனையால் அழியும் தன்மை உடைத்தாயின் இம்முறை முடியாததாய் அமைந்துவிடுகிறது.

நடைமுறையில் எடுக்கப்படும் பெரும்பாலான மாதிரிகள் வரையறுக்கப்பட்ட முழுமைத் தொகுதியிலிருந்து (finite population) எடுத்த பொருளைச் சேர்க்காமல் எடுக்கும் மாதிரிகளாகவே (sampling without replacement) உள்ளன. உறுப்புகளை எடுக்கும் சோதனை தொடர்ந்து நடைபெறும்போது முழுமைத் தொகுதியில் உள்ள உறுப்புகள் குறைவதோடு உறுப்புகள் எடுப்பதற்கான நிகழ்திறனும் மாறுபடுகிறது, முழுமைத் தொகுதியுடன் ஒப்பிடும்போது மாதிரியின் அளவு சிறியதாய இருப்பின், இது வரையறுக்கப்படாத முழுமையிலிருந்து எடுக்கப்பட்ட மாதிரிக்குச் சமமாகி விடுகிறது. மாதிரியிலுள்ள ஒரு பண்பிற்கான விகிதசமம் முழுமைத் தொகுதியிலுள்ள அப்பண்பிற்கான விகிதசமத்தைப் பிரதிபலிக்காத தன்மை குறிப்பிடும்படி அமைந்திருப்பின், இதற்கான திருத்தம் ஏற்படுவது தேவையானதாகும்.

பகுப்பு முறை இலையிலா மாதிரி (Stratified Random Sample)

மாதிரி, முழுமைத் தொகுதியின் பிரதிநிதித்துவத் தன்மையைக் (representative) கொண்டதாயும், மாதிரி உறுப்புகள் சமவாய்ப்பைக் (equal probability) கொண்டு தேர்ந்தெடுக்கப்பட்டதாயும் இருக்கவேண்டிய நிலையைக் கூறினோம். முழுமைத் தொகுதியின் தன்மை, சமவாய்ப்புடைய உறுப்புகளைக்கொண்ட மாதிரிகளை எடுத்தபோதிலும் முழுமைத் தொகுதிக்கான பிரதிநிதித்துவத் தன்மையைப் பெறுமல இருப்பதைக் காணமுடிகிறது. மாதிரியிலிருந்து பெறப்படும் முடிவுகள் முழுமைத் தொகுதிக்கும் பொருந்த வேண்டி உள்ளதால சமவாய்ப்புப் பெற்று அமையும் உறுப்புகள் முழுமைத் தொகுதியைப் பிரதிபலிப்பதாகவும் அமையவேண்டிய தேவை ஏற்படுகிறது. முழுமைத் தொகுதி ஒரே தன்மைத்தான (homogeneous) உறுப்புகளைக் கொண்டமையாவிட்டால், அவற்றின் மாறுபாட்டின் தன்மைக்கேற்ப நமது மாதிரியில் உள்ள உறுப்புகள் சமவாய்ப்புக் கொடுத்து தேர்ந்தெடுக்கப்பட்டபோதிலும், முழுமைத் தொகுதி

ஒரே தன்மைத்தான (homogeneous) உறுப்புகளைக் கொண்டமையவிட்டால், அவற்றின் மாறுபாட்டின் தன்மைக்கேற்ப நமது மாதிரியில் உள்ள உறுப்புகள் சமவாய்ப்புக் கொடுத்துத் தேர்ந்தெடுக்கப்பட்ட போதிலும், முழுமைத் தொகுதிக்கான உண்மையான பிரதிநிதியாக (true representative) அமையாது. மேற்சொன்ன 30 வயல்களும் ஒரே பரப்புள்ளதாக அமையாவிட்டால், நம்மால் தேர்ந்தெடுக்கப்பட்ட வயல்கள் எல்லாம் பெரியனவாகவோ அல்லது சிறியனவாகவோ அமைய இடமிருக்கிறது. சமவாய்ப்புக் கொடுக்கும்போது இப்படி அமையாதென நம்பினாலும், இப்படி அமைவதற்குள்ள வாய்ப்பைத் தள்ளிவிட முடியாது. இயைபிலா மாதிரியாக அமையப்பெற்றாலும், ஒரு பிரதிநிதித்துவத் தன்மையைப் பெறுமல் இருப்பதை உணர்கிறோம். இக்குறையைப் பகுப்புமுறை இயைபிலா மாதிரி (stratified random sampling) எடுக்கும் தன்மையால் குறைக்கலாம்.

முழுமைத் தொகுதியிலுள்ள உறுப்புகள் ஒரே தன்மைத்தாய் அமையாமல் (hetrogeneous) இருக்குமாயின் அவற்றின் மாறுபாட்டின் தன்மைக்கேற்ப முழுமைத் தொகுதியை ஓரளவிற்கு ஒரே தன்மைத்தான உறுப்புகள் (homogeneous) கொண்ட பகுதிகளாகப் பிரிக்கவேண்டும். இம் முறைப்படி பிரிக்கும் தன்மைக்குப் பகுப்புமுறை (stratification) என்றும், பகுத்த பகுதிகளுக்குப் பகுதி (stratum) என்றும், இப் பகுப்பிலிருந்து எடுக்கப்படும் மாதிரிக்குப் பகுப்புமுறை இயைபிலா மாதிரி (stratified random sampling) என்றும் பெயர். இதன் அடிப்படைத் தத்துவம் மிக எளிதானது. முழுமைத் தொகுதியைப் பல பகுதிகளாகப் பிரித்து, ஒவ்வொரு பகுதியிலிருந்தும் முன்பு கண்ட இயைபிலா மாதிரி எடுக்கும் முறையைக் கொண்டு எடுத்துச் சேர்த்த மாதிரி பகுப்புமுறை இயைபிலா மாதிரி எனப்படும். வயல்களின் பரப்பிற்குத் தக்கவாறு 12, 12, 6 என்ற மூன்று பகுதிகளாகக் கொண்டு 2, 2, 1 என்ற முறைப்படி 5 வயல்களைத் தேர்ந்தெடுக்கலாம். முதல் வகுப்பிற்கு இரண்டும், இரண்டாவது பகுப்பிற்கு இரண்டும், மூன்றாவது பகுப்பிற்கு ஒன்றுமாக ஐந்து வயல்களைச் சம வாய்ப்பினைக் கொடுத்துத் தேர்ந்தெடுத்து ஒரு பிரதிநிதித்துவ மாதிரியாகக் கொள்கிறோம். இம் முறை பகுதிகளின் வேறுபாட்டைக் (variation) குறைப்பதோடு மதிப்பீடுகளின் (estimation) தன்மை உயர்ந்ததாய் அமையவும் ஏதுவாய் உள்ளது. பகுப்புமுறை அடிப்படை மாறுபாட்டிற்கான எந்த அளவாகவேனும் இருக்கலாம். இடம், பரப்பு, வயது, உருவம் ஆகிய எதைக் கொண்டாயினும் பகுக்கப்படலாம். பண்பின் அடிப்படையில் அல்லது அளவின் அடிப்படையில்கூடப் பிரிக்கப்

படலாம். இரண்டு அல்லது அதற்கு மேற்பட்ட பண்புகள் மாறு பாட்டிற்குக் காரணமாயின், இவையனைத்தும் பகுப்பிற்குக் காரணமாய் அமைந்து பகுதிகளை அதிகப்படுத்தும். பகுப்பிற்குத் தக்கவாறு மூன்று பகுதிகளாகக்கொண்ட மூன்று பிரிவுகளையும் பயிர்களின் விதத்திற்குத் (variety) தக்கவாறு மேலும் 3 பிரிவு களாகப் பிரித்தால் நமக்குக் கிடைப்பது 3×3 பகுதிகளாகும். இந்த ஒவ்வொரு பகுதியிலிருந்தும், இயைபிலா ராண்டம் கூறெடுத்தலின் தன்மையைக்கொண்டு மாதிரியின் உறுப்புகளைத் தேர்ந்தெடுக்கவேண்டும். முழுமைத் தொகுதியைப் பகுப்பதற் கான காரணங்களும், பகுப்பு முறையும் தக்க காரணங்கட்குட் படாத நிலையில், பகுபடாத நிலையில் எடுக்கப்படும் இயைபிலா மாதிரிகளின் முடிவுகளைவிட மோசமான முடிவுகளைக் கொடுக்கும் என்பதை உணர்ந்து செயல்படவேண்டும்,

மாதிரி உறுப்பு (Sampling Unit)

மாதிரி எடுப்பதற்கு எவை அடிப்படை உறுப்புகளாக அமைகின்றனவோ அவையே மாதிரி உறுப்புகள் (sampling units) எனக் கொள்ளப்படும். சென்னை நகர மக்கள்தொகையைக் கணிக்கும்போது முழுமைத் தொகுதியின் உறுப்புகள் தனிப்பட்ட நபர்களாகவும் மாதிரியின் உறுப்புகள் தனிப்பட்ட நபர்களா கவோ வீடாகவோ, அல்லது மாநகராட்சியின் ஒரு வட்ட மாகவோ அமையலாம். முழுமைத் தொகுதி உறுப்புகளே மாதிரி யின் உறுப்புகளாக இருக்குமாயினும், மாதிரியின் உறுப்பு எனக் குறிப்பிடப்படுவது தனிப்பட்ட ஓர் உறுப்பாகவோ அல்லது பல பண்புகளின் கூட்டாகவோ கூறெடுத்தலுக்கு அடிப்படையாக அமையும் உறுப்பினையேயாகும். பள்ளி மாணவர்கள் தம்கூட்டுச் சராசரி உயரம் காண எடுக்கப்படும் ஒரு மாதிரியில், முழுமைத் தொகுதியின் உறுப்புகள் பள்ளி மாணவர்களின் உயரங்களாகவும், மாதிரியின் உறுப்புப் பள்ளிகளாகவோ அல்லது பள்ளியிலுள்ள வகுப்புகளாகவோ அமைவதைக் காணலாம், மாதிரியின் உறுப்புகள் இறுதியில் மாணவர்களாக அமைந்தாலும், பள்ளி களோ அல்லது வகுப்புகளோ மாதிரி எடுப்பதற்கு அடிப்படையாக அமைவதால் பள்ளியை அல்லது வகுப்பினை மாதிரி உறுப்பு (sampling unit) எனச் சொல்கிறோம்.

ஒரே தன்மைத்தான உறுப்புகள்கொண்ட இடத்திலும், அப்படி அமையாத இடத்திலும் இயைபிலா மாதிரி எடுப்பிற் கான சில அடிப்படைக் கருத்துகளைக் கண்டோம். இம் மாதிரி எவ்வளவு உறுப்புகளைக் கொண்டதாய் இருக்கவேண்டுமென்ப தற்கான சில பண்புகளைக் காண்போம்.

மாதிரி அளவு (Sample size)

மாதிரியின் உருவளவு அல்லது அளவு (sample size) என்பது, முழுமைத் தொகுதியிலிருந்து ராண்டமாக எடுக்கப்பட்ட மாதிரியிலுள்ள உறுப்புகளின் எண்ணிக்கையைக் குறிக்கும்.

மாதிரியின் அளவு, முழுமைத் தொகுதியிலுள்ள பிரச்சினைக்குரிய பண்பின் மாறுபாட்டை (variation) அடிப்படையாகக் கொண்டமைவதை முதலாவது பண்பாகக் கொள்ளலாம். அதிக மாறுபாடு உள்ள முழுமைத் தொகுதிக்கு பெரிய உருவளவு (large sample) மாதிரியாகவும், குறைந்த மாறுபாடுள்ள முழுமைத் தொகுதிக்கு குறைந்த அளவுள்ள (small samples) மாதிரியாகவும் இருக்கவேண்டும். ஒத்த தன்மைத்தான (perfectly homogeneous) முழுமைத் தொகுதியிலிருந்து எடுக்கப்படும் ஒரு பொருளே முழுமையின் பண்பைப் பிரதிபலிப்பதாக அமைந்துவிடும். நமது உடம்பிலிருந்து எடுக்கப்படும் ஒரு சொட்டு ரத்தம், நமது ரத்தம் முழுவதையும் பிரதிபலிப்பதாகக் கொள்கிறோம். ஒத்த உயரமுடைய மாணவர் குழுவிலிருந்து எடுக்கப்படும் ஒரு மாணவன் அம் முழுமைத் தொகுதிக்குமான பண்பைப் (உயரத்தை) பிரதிபலிப்பவனாக அமைகிறான். மாறுபாடு நிறைந்த முழுமைத் தொகுதியிலிருந்து ராண்டமாக எடுக்கப்படும் அதிக அளவுள்ள மாதிரியே முழுமைத் தொகுதியின் பண்பைப் பிரதிபலிப்பதாக அமையும். அடிப்படைப் பணியில் மாறுபாடு நிறைய இருப்பின் பெரிய அளவுள்ள மாதிரிகள் எடுப்பது அவசியம்.

நமது சோதனைக்குட்படும் முழுமைத் தொகுதியின் பண்பு அல்லது குணம் மிக அரிதாக (rare) அமையப் பெற்றிருப்பின் இப் பண்பை அடிப்படையாகக்கொண்டு எடுக்கப்படும் மாதிரியின் அளவு அதிகமாக அமைய வேண்டும். மாதிரியின் அளவைக் கணக்கிட உதவி செய்யும் இரண்டாவது யுத்தியாகக் கொள்ளலாம். விதைப் பொட்டலமொன்றிலிருந்து கிடைக்கும் பூக்களின் நிறம் வெள்ளை மற்றும் சிவப்பாக இருப்பதற்கான விகித சமம் $1:9$ எனக் கொள்வோம். மாதிரிப் பருமன் ஐந்தாக உள்ள மாதிரியில் சிவப்புப் பூக்கள் கிடைக்காதிருப்பதற்கான நிகழ்திறனைக் கணக்கிடுவோம். வெள்ளைப் பூ கிடைக்காமலிருப்பதற்கான நிகழ்திறன் $(0{\cdot}9)^5 = 0{\cdot}59$ சராசரியாக ஐந்து உறுப்புகள் கொண்ட ஐந்து மாதிரிகளில் மூன்று மாதிரிகளில் வெள்ளைப் பூ கிடைக்காமலிருக்கும். சிவப்புப் பூ கிடைக்காமலிருப்பதற்கான நிகழ்திறன் $({\cdot}1)^5 = {\cdot}00001$. மாதிரிப் பருமன் ஐந்தாக உள்ள மாதிரிகளில் சிவப்புப் பூ கிடைக்காமலிருப்பதற்கான வாய்ப்பு

மிகக் குறைவாக உள்ளது. நமது மாதிரி முழுமைத் தொகுதியைப் பிரதிபலிப்பதாக அமைய, மாதிரியின் அளவை அதிகரித்து அமைக்க வேண்டும். அரிதாக நிகழும் (are events) பண்புகளைக் கணக்கிடும்போது மாதிரியின் அளவு அதிகப்படுத்தப்பட வேண்டுமென்பதை இது உணர்த்துகிறது.

மாதிரியின் அளவைத் திட்டமிட உதவும் மூன்றாவது நெறியாக 'இரு தொகுதிகளின் (group) பண்புகளுக்கான வித்தியாசம் மிகவும் குறைந்துள்ள இடத்தில் பெரிய அளவுள்ள (large sample) மாதிரிகளே முழுமைத் தொகுதியிலுள்ள பண்பின் தன்மையையும், வித்தியாச அளவையையும் குறிப்பிடும்படி அமையும்' என்பதைக் கூறலாம். மாதிரியிலிருந்து கிடைக்கக் கூடிய முடிவு, முழுமைத் தொகுதியிலுள்ள 80 மற்றும் 20 என்ற விகிதத்தில் அமைந்த இரு பண்புகளைப் பிரதிபலிப்பாக இருப்பின், சிறிய அளவு மாதிரியே போதுமானதாக உள்ளது. ஆனால், முழுமைத் தொகுதியிலுள்ள பண்புகளின் விகிதம் 52 : 48 என அமையும்போது பெரிய அளவுள்ள மாதிரிகளை (large samples) எடுக்க வேண்டியவர்களாக உள்ளோம். முதல் தன்மையான இடத்தில் (80 : 20) சில நூறு உறுப்புகளே போதுமானதாகவும், இரண்டாவதாகக் குறிப்பிட்ட சூழலில் (52 : 48) சில ஆயிரம் உறுப்புகள் தேவை என்பதையும் உணர்தல் வேண்டும்.

உறுப்புகள் எடுப்பதற்கான வாயப்புச் சமமாக இருக்கும் ராண்டம் மாதிரிகளை எடுக்கும்போது, மாதிரியின் அளவு அதிகப்பட மாதிரியிலிருந்து பெறப்பட்ட மதிப்பீடுகள், முடிவுகளின் நம்பகத் தன்மையும் அதிகமாக்கப்படுகிறது. மாதிரியின் அளவை அதிகப்படுத்துவதால்மட்டும் பிறழ்ச்சியற்ற (unbiased) மதிப்பீடுகளையும், முடிவுகளையும் செய்யலாம் என்ற முடிவிற்கு வர முடியாது. பிரதிநிதித்துவத் தன்மை இல்லாத (representative) உறுப்புகளைச் சேர்ப்பதால் விளைவுகள் விபரீதமாக முடியும். சேர்க்கப்படும் உறுப்புகள் அனைத்தும் சமவாய்ப்பினைப் பெற்றதாய், முழுமைத் தொகுதியின் பண்பிற்குப் பிரதிநிதித்துவக் கூறாக அமையும்போது நமது முடிவுகளும், மதிப்பீடுகளும் பிறழ்ச்சியற்று அமைவதற்கான நிகழ்திறன் அதிகமாக்கப்படுகிறது.

முழுமைத் தொகுதியில் உள்ள ஒவ்வோர் உறுப்பும் மாதிரியில் சேர்க்கப்படுவதற்குச் சமவாய்ப்பினைப் பெறும்போது, மாதிரியின் பருமன் அதிகரிக்கப்பட்டால், மாதிரியின் அளவைகள் (statistic) முழுமைத் தொகுதியின் அளவைகளைப் (parameter)

பிரதிபலிக்கும் தன்மை அதிகப்படுத்தப்படுகிறது. அடிக்கடி நிகழும் தன்மைத்தான உறுப்புகள் மிக விரைவாகத் தேர்ந்தெடுக்கப்படும் தன்மையாயும், அரிதாக நிகழும் நிகழ்ச்சிகள் மாதிரியின் அளவு அதிகப்படும்போது நிகழும் தன்மைத்தாயும் அமைகின்றன. மாதிரிப் பருமன் அதிகமான இயைபிலா மாதிரிகளில்தான் அரிதான நிகழ்ச்சிகளையும், குறைந்த வித்தியாசம் உடைய இரு பண்புகளையும் காணும் தன்மை அதிகமாக்கப்படுகிறது.

மாதிரியின் திறம் (efficiency) என்பது n_1 உறுப்புகள் கொண்ட மாதிரியின் ஓர் அளவைக்கான வேறுபாட்டிற்கும் (sampling variance), அதே அளவைக்கான n_2 உறுப்புகள் கொண்ட மற்றொரு மாதிரியின் வேறுபாட்டிற்கும் உள்ள விகித சமத்தைக் குறிக்கும். இரண்டாவது மாதிரியின் மாறுபாட்டளவு மிகக் குறைந்ததாய் அமைய அமைய விகித சமத்தின் அளவு பெரிதாகிக்கொண்டே செல்லும். இச் சூழலில் இரண்டாவது மாதிரியின் திறம் அதிகரிக்கப்படுகிறது. எந்த மாதிரியின் ஓர் அளவைக்கான மாறுபாட்டளவு குறைந்ததாக உள்ளதோ அம் மாதிரியின் திறம் அதிகமானது (இரண்டையும் ஒப்பிடும் போது) எனக் கொள்கிறோம்.

திறமுடைய இயைபிலா மாதிரிகள் எடுப்பதற்கான சில அடிப்படைக் கருத்துகளைக் கண்டோம். புள்ளியியல் துறையில் மாதிரியின் தன்மைக்கான படிப்பு ஒரு தனிப்பட்ட துறையாக வளர்ந்துவிட்டது. இதற்கான முழுக்கருத்தினையும் கூற முற்படுவது முடியாததாயும் நமது நோக்கத்திற்குத் தேவையற்றதாயும் உள்ளது. கூறுகள் எடுக்கும்போது மேற்சொன்ன கருத்துகளை மனத்திற் கொள்வது பயனுடைத்தாகும். கூறின் அளவுகள் எப்படிக் கணக்கிடப்படுகிறது என்பதைச் சிறப்புக் காணும் சோதனைப் (test of significance) பகுதியில் காண உள்ளோம். சிறப்புக் காணும் சோதனைகளின் அடிப்படைக் கருத்துகளாக மாதிரியின் பரவல் மற்றும் திட்டப்பிழைபற்றிப் பார்ப்போம்

மாதிரியின் பரவல் (Sampling Distribution)

'N' உறுப்புகள்கொண்ட முழுமைத் தொகுதியிலிருந்து n உறுப்புகள்கொண்ட k மாதிரிகள் எடுப்பதாகக் கொள்வோம். மாதிரி ஒவ்வொன்றிற்கும் கூட்டுச் சராசரி ($\bar{x}$), திட்ட விலக்கம் s) உடன் தொடர்புகெழு (r_{xy}) ஆகியவற்றைக் கணக்கிட்டு ($\bar{x}_1$, $\bar{x}_2$ ......$\bar{x}_k$ S_1 S_2...... S_k, r_1 r_2......r_k எனக் குறிப்போம்.

மாதிரியிலிருந்து இது போன்று எந்த அளவையையும் கணக்கிட்டுக் கொள்ளலாம். $\overline{x_1}, \overline{x_2}, \ldots \overline{x_k}$ என்ற கூட்டுச் சராசரிகளை மட்டும் எடுத்துக்கொள்வோம். இதற்கான முடிவுகள் மற்ற அளவைகளுக்கும் பொருந்தும். மாதிரியை எடுத்து அதற்கான அளவைகளைக் கணக்கிடுவது, முழுமைத் தொகுதியின் அப் பண்பைக் காணவே என்பதை முன்பு குறிப்பிட்டோம். $\overline{x}_1, \overline{x}_2, \ldots \overline{x}_k$ ஆகியவற்றை ஒரு நிகழ்வெண் பரவலாக அமைப்போம். மாதிரியின் அளவு (size) அதிகமாக்கப்படும் போது, இந்த நிகழ்வெண் பரவலுக்குக் கிடைக்கும் வளை கோட்டிற்கான நிகழ்வெண் சார்வே (frequency function) மாதிரியின் நிகழ்வெண் பரவல் எனப்படும். இந் நிகழ்வெண் பரவலுக்கான அளவையின் பெயராலேயே இந் நிகழ்வெண் பரவல் சுட்டப்படும். கூட்டுச் சராசரிக்கான மாதிரியின் பரவல், மாறுபாட்டளவைக்கான மாதிரியின் பரவல், உடன்தொடர்பு கெழுவிற்கான மாதிரியின் பரவல் என மாதிரியின் அளவைகளின் தன்மைக்கேற்பப் பெயர்பெறும்.

தனித்த அளவுகள் இயல்நிலைப் பரவலாக அமையாவிடத்தும், மாதிரியின் அளவைகளுக்கான பரவல்கள் இயல்நிலைப் பரவல்களாகவோ அல்லது இயல்நிலைப் பரவலை ஒத்தோ அமையும் சிறப்புக் குறிப்பிடத்தக்கது. சிறிய அளவுள்ள (small sample) மாதிரிகள் உள்ள இடத்தும், அவற்றின் அளவுகளுக்கான பரவல்கள் ஒரு தனித்த நிகழ்திற அமைப்பைப் பெறுவது குறிப்பிடத்தக்கதாகும்.

திட்டப்பிழை (Standard Error)

மாதிரியின் பரவலுக்கான மாறி (variate), முழுமைத் தொகுதியின் ஒரு பண்பின் மதிப்பீட்டிற்கான பல்வேறுபட்ட அளவுகளைக் குறிக்கும். மாதிரியின் பரவல் ஒன்றிற்கான திட்ட விலக்கமே (standard deviation) திட்டப்பிழை (standard error) எனப்படும். திட்டவிலக்கம் மாறிகளின் தனிப்பட்ட அளவுகளிலிருந்து கணக்கிடப்படும். $k_1, k_2 \ldots k$ என்ற k மாதிரிகளுக்கும் தனித்தனியாகக் கூட்டுச் சராசரி கணக்கிடப்படும். ஆனால், திட்டப்பிழை என்பது இக் கூட்டுச் சராசரிகளை நிகழ்வெண் பரவலாக அமைத்து அதற்குக் கணக்கிடப்படும் திட்ட விலக்கமாகும். திட்ட விலக்கத்தின் உறுப்புகள் தனித்த அளவுகளாகவும், திட்டப்பிழையின் உறுப்புகள் ஒவ்வொன்றும் மாதிரியின் பல உறுப்புகளைக்கொண்டு கணக்கிடப்பட்ட அளவைகளாகவும் இருப்பதை உணர்த்தல் வேண்டும். மாதிரிகளால் மதிப்பீடு செய்யப்பட்ட முழுமைத் தொகுதியின் ஓர்

அளவை 100 மாதிரிகளில் 95 மாதிரிகளின் மதிப்பு, அம் மதிப்பீட்டிலிருந்து இருபுறமும் 1·96 பங்கு திட்டப்பிழை இடைவெளியில் அமைவதற்கான நம்பிக்கையை நாம் ஏற்கின்றோம். மாதிரியின் அளவைகளின் தன்மைக்குக் கட்டுப்பட்டு மாதிரிப் பரவல்கள் அமைகின்றன. எந்த அளவையின் மாதிரிப் பரவலுக்குத் திட்ட விலக்கம் கணக்கிடப்படுகிறதோ அந்த அளவையின் பெயராலேயே திட்டப்பிழையும் குறிக்கப்படும். கூட்டுச் சராசரியின் திட்டப்பிழை (standard error of mean), திட்டவிலக்கத்தின் திட்டப்பிழை (standard error of standard deviation), உடன்தொடர்பு கெழுத் திட்டப்பிழை (standard error of the correlation co-efficient) எனத் திட்டப்பிழைகளைக் குறிப்பிடுகிறோம். கணக்கியல் முறைப்படி கணக்கிடப்பட்டுள்ள அடியிற் காணும் திட்டப்பிழைகளை மனத்திற் கொள்வது சிறப்புக்காணும் சோதனைகள் செய்வதற்கு ஏதுவாக இருக்கும்.

மாதிரி எடுத்தல் (பண்புகளின் அடிப்படையானது)
(Sampling of Attributes)

அளவை	திட்டப்பிழை
வெற்றிகளின் எண்ணிக்கை (np)	$\sqrt{npq}$
வெற்றிகளின் விகித சமம் p	$\sqrt{pq/n}$
விகிதசமங்களின் வித்தியாசம் $p_1 - p_2$	$\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ $p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$
விகிதசமங்களின் வித்தியாசம் $p_1 - p_2$	$\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$

மாதிரி எடுத்தல் (அளவைகளின் அடிப்படையானது)
(Sampling of Variables)

அளவை	திட்டப்பிழை
கூட்டுச் சராசரி x	$\sigma / \sqrt{n}$
கூட்டுச் சராசரிகளுக்கான வித்தியாசம் $\bar{x}_1 - \bar{x}_2$	(a) $\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n}}$ (b) $\sqrt{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ $= \sigma \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$
திட்ட விலக்கம் σ	$\sigma / \sqrt{2n}$
மாறுபாட்டளவை σ^2	$\sigma^2 \sqrt{2/n}$
உடன்தொடர்புகெழு $= rxy$	$(1-r^2) / \sqrt{n}$
மாறிகளின் உடன்தொடர்பு கெழு (Regression co-efficient)	$\frac{\sigma x \sqrt{1-r^2}}{\sigma y \sqrt{n}}$

பயிற்சி :

1. நமது மாநிலத்தில் ஓர் ஆண்டில் கிடைக்கும் கேழ்வரகு விளைச்சல் மதிப்பீடு (estimation) செய்யும் முயற்சிக்கான ஒரு மாதிரித் திட்டத்தை (sampling plan) அமை. எடுக்கப்படும் மாதிரியின் தன்மைக்கான காரணங்களைக் கூறு.

2. சென்னை நகரத்தின் வீட்டு நெருக்கடியை over crowding) உணர எடுக்கப்படும் மாதிரியின் திட்டம்பற்றி விரிவாக்கு.

3. புள்ளியியல் முறைப்படி குறிப்பிடப்படும் முழுமைத் தொகுதியின் தன்மை யாது ? புள்ளியியல் துறையில் குறிப்பிடப்

படும் முழுமைத் தொகுதியின் தன்மைகளுக்கு இரண்டிரண்டு உதாரணம் கொடு.

4. எடுத்த பொருளைச் சேர்த்தபின், எடுத்த பொருளைச் சேர்க்காமல் எடுக்கப்படும் மாதிரிகளின் தன்மை யாது? வரையறுக்கப்பட்ட முழுமைத் தொகுதியில் இவற்றுக்கான மாறுபாட்டை ஓர் எடுத்துக்காட்டால் விளக்கு.

5. பகுத்தெடுக்கும் இயைபிலா மாதிரியின் நன்மை யாது? மாறுபட்ட தன்மை உள்ள முழுமைத் தொகுதியில் பகுத்தெடுக்கும் முறை செயல்படாவிடத்து ஏற்படும் விளைவுகளை ஒரு சிறந்த எடுத்துக்காட்டால் விளக்கு.

6. மாதிரியின் அளவு (sample size) பற்றி விவரி.

10. புள்ளியியல் அனுமானம்

(Statistical Inference)

மதிப்பீடு மற்றும் கருதுகோட் சோதனைகள்

(Estimation and testing of Hypothesis)

நாம் செய்யும் சோதனைகளுக்கான உருப்படிவங்கள் (statistical models) அமைப்பதே புள்ளியியலாரின் முக்கிய வேலையாக அமைந்துள்ளது. இவ் வுருப்படிவங்கள், சோதனைகளின் தன்மையைப் படிப்பதற்கும், இம் முறையான சோதனைகளைத் தொடர்வதால் நிகழக்கூடிய எதிர்கால முடிவுகளின் தன்மையை உணரவும் உதவியாக இருக்கிறது. இவ் வடிவங்களின் உதவியின்றி, புள்ளியியல் முறைப்படி சோதனைகளையும், மதிப்பீடுகளையும் செய்வது கடினமானதாகவோ அல்லது முடியாததாகவோ அமைந்துள்ளது.

உருப்படிவங்கள் அமைப்பதும், முழுமைத் தொகுதிக்கான அளவைகளை மதிப்பீடுவதும், மதிப்பிட்ட அளவைகளுக்கு நாம் கொடுக்கும் நம்பிக்கையைக் கணக்கிடுவதும் புள்ளியியல் அனுமானம் (statistical inference) ஆகும். புள்ளியியல் அனுமானம் பெற்றறிமுறை (deductive) யாகவும், பட்டறிமுறை (inductive) யாகவும் அமையலாம். பெற்றறி முறையில் முழுமைத் தொகுதிக்கான பண்புகள் முழுதும் தெரியப் பெற்றவர்களாக, அம் முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட மாதிரியின் அளவுகளைக் கொண்டு முழுமைத் தொகுதி அளவுகளை (parameter) அனுமானம் செய்கிறோம். நாணயத்தை வீசுவதையும், சீட்டுக் கட்டொன்றிலிருந்து எடுக்கப்படும் சீட்டுகளையும் முழுமைத் தொகுதியின் உறுப்புகளாகக் கொண்டால் இதற்கான உருப்படிவம் ஈருறுப்புப் பரவலாக (binomial distribution) அமையும் என்பதை முன்பு கண்டோம். இது பெற்றறி முறையாகும். நமது

ஆய்விற்குரிய பெரும்பாலான சிக்கல்கள் முழுமைத் தொகுதிக்கான பண்புகள் தெரியாமலிருக்கும் பட்டறி முறை அனுமானமாக அமைந்துள்ளன. இம் முழுமைத் தொகுதியிலிருந்து எடுக்கப்படும் மாதிரிகளிலிருந்து மதிப்பீடு மற்றும் கருதுகோள் சோதனைகள் செய்து முழுமைத் தொகுதிக்கான பண்புகள் இம் மாதிரி இருக்கலாம் என அனுமானிக்கப்படுகிறது. நமது பிரச்சினைக்குரிய முழுமைத் தொகுதியை நாம் அறியாத காரண மட்டுமன்றி முழுமைத் தொகுதிக்கான அளவுகள் மாறும் தன்மைத்தாய் அமையும்போது இப் பட்டறி முறையின் உபயோகம் மேலும் வலுப்பெறுகிறது.

நிகழ்ச்சிக்கான உருவமைப்பை அனுமானிப்பதைக் காட்டிலும், நிகழ்ச்சியோடு ஒத்த ராண்டம் மாறிகளுக்கான (random variable) உருவமைப்பைத் திட்டமிடுவதே நமது நோக்கமாகும். x என்பது 100 விதைகள் கொண்ட பொட்டலத்தில் உள்ள பழுதான விதைகளின் எண்ணிக்கையைக் குறிப்பதாகக் கொள்வோம். எடுக்கப்படும் ஒரு பொட்டலத்தில் எவ்வளவு விதைகள் பழுதானவை என்பதைத் தெரிய உதவி செய்யும் உருவமைப்பைக் காணுவதைவிட, x மாறியின் பல்வேறுபட்ட மதிப்புகளைக் காட்டும் உருவமைப்பைத் திட்டமிடுவதே நமது நாட்டமாகும். பெரும்பாலான உருவமைப்புகள் ராண்டம் மாறிகளுக்கான நிகழ்வெண் பரவலைக் குறிக்கும். இயைபிலா மாறிகளுக்கான நிகழ்வெண் பரவலைப் பற்றிய அனுமானங்களே புள்ளியியல் அனுமானங்கள் எனப்படும்.

புள்ளியியல் அனுமானமாவது நிகழ்திறன் உதவியால் எடுக்கப்படும் முடிவுகளாகும். புள்ளியியல் அனுமானங்கள் பெரும்பாலும் 'மதிப்பீடு' செய்வதாகவோ அல்லது 'கருதுகோட் சோதனைகள்' செய்வதாகவோ அமைந்துள்ளன.

மதிப்பீடு (Extimation)

மதிப்பீடு என்பது முழுமைத் தொகுதி அளவை (parameter) ஒன்றிற்கான மதிப்பை மாதிரி அளவுகளின் (sample observations) உதவியால் மதிப்பிடும் முறையைக் குறிக்கும். மதிப்பீடு செய்யப்படவேண்டிய (estimator) அளவைகளுக்கான மாதிரி அளவுகள் மதிப்பீட்டளவைகள் (estimates) எனவும், மதிப்பீடு செய்யும் முறை மதிப்பீட்டு முறை (estimation) எனவும் சொல்லப்படும். முழுமைத் தொகுதியின் அளவுகள் அனைத்தும் கொண்டு பெறப்படும் அளவைகளுக்கு, 'முழுமைத் தொகுதி அளவை'

(population parameter) எனவும்), மாதிரியின் அளவைகளைப் பயன்படுத்திப் பெறப்படும் அளவைகள் மாதிரி அளவை (sample statistic) எனவும் குறிக்கப்படும். முழுமைத் தொகுதி அளவைகளை கிரேக்க எழுத்துகளாலும், மாதிரி அளவைகளை ஆங்கில எழுத்துகளாலும் குறிப்பது வழக்கம். கூட்டுச் சராசரியை 'μ' எனவும், திட்ட விலக்கத்தை 'σ' எனவும், உடன் தொடர்பு கெழுவை 'ρ' எனவும் குறிப்பிட்டால் முழுமைத் தொகுதி அளவைகளையும், கூட்டுச் சராசரியை '$\bar{x}$' எனவும், திட்டவிலக்கத்தை 'S' எனவும், உடன் தொடர்பு கெழுவை 'r' எனவும் குறிப்பிட்டால் மாதிரி அளவைகளையும் குறிக்கும்.

ஒரு கல்லூரி மாணவர்களில் 10 மாணவர்களை எடுத்து அவர்தம் உயரங்களைக் கணக்கிடுவோம். 67·2″, 64·7″, 69·1″, 63·8″, 68·5″, 70·0″, 67·7″, 66·3″, 68·4″, 65·2″ எனப் பத்து மாணவர்களது உயரங்களும் குறிக்கப்பட்டுள்ளன. இவ் வளவுகளுக்கான கூட்டுச் சராசரி ($\bar{x}$) 67·1″ என அமைகிறது. பத்து மாணவர்கள் கொண்ட மாதிரி, கல்லூரியில் உள்ள ஒவ்வொரு மாணவருக்கும் சமவாய்ப்புக் (equal probability) கொடுத்து எடுக்கப்பட்ட இயைபிலா மாதிரியாக (random sample) அமையுமிடத்து, 67·1 என்ற மாதிரியின் கூட்டுச் சராசரியே கல்லூரி மாணவர்களின் கூட்டுச் சராசரி உயரத்திற்கு ஒரு பிறழ்ச்சியற்ற மதிப்பீடாகக் (unbiased estimate) கொள்ளப்படும்.

முழுமைத் தொகுதி அளவுகளையும் குறிப்பிட்டு, அதிலிருந்து எடுக்கப்படும் இயைபிலா மாதிரி, அம் முழுமைத் தொகுதி அளவையைச் (parameter) செம்மையாக மதிப்பீடு செய்யும் தன்மையைப் பார்ப்போம்.

முழுமைத் தொகுதியின் அளவுகள் 1, 2, 3, 4, 5, 6 எனக் கொள்வோம். இவற்றில் இரண்டும், ஐந்தும் நமது பண்புகளைக் குறிப்பதாகக் கொள்வோம். இதற்கான நிகழ்திறன் $\frac{2}{6} : \frac{1}{3}$ என அமைகிறது. இதனைக் கணக்கிட இதிலிருந்து எடுக்கப்படும் மாதிரிகளை உபயோகப்படுத்தும் முறையைக் காணலாம். ஆறு உறுப்புகளிலிருந்து நான்கு உறுப்புகள் கொண்ட மாதிரிகள் அனைத்தையும் அமைத்து, அவற்றில் நமது பண்பிற்கான உறுப்புகள் இருக்கும் எண்ணிக்கையையும் அதற்கான நிகழ்திறனையும் கணக்கிடுவோம்.

மாதிரியின் அளவுகள்	குறிப்பிட்ட பண்பின் அளவு	மாதிரியின் அளவுகள்	குறிப்பிட்ட பண்பின் அளவு	மாதிரியின் அளவுகள்	குறிப்பிட்ட பண்பின் அளவு
1 2 3 4	1	1 2 5 6	2	2 3 4 5	2
1 2 3 5	2	1 3 4 5	1	2 3 4 6	1
1 2 3 6	1	1 3 4 6	0	2 3 5 6	2
1 2 4 5	2	1 3 5 6	1	2 4 5 6	2
1 2 4 6	1	1 4 5 6	1	3 4 5 6	1

இவ் வளவுகளில் குறிப்பிட்ட பண்பு அமைவதற்கான நிகழ்திற நிகழ்வெண் பரவலை அமைப்போம்.

குறிப்பிட்ட பண்பு அமையும் நிலை	நிகழ்திறன்	மாதிரிகளின் எண்ணிக்கை
0	0	1
1	$\frac{1}{4}$	8
2	$\frac{1}{2}$	6
		15

இந் நிகழ்வெண் பரவலின் கூட்டுச் சராசரி $= \dfrac{0 \times 1 + \frac{1}{4} \times 8 + \frac{1}{2} \times 6}{15}$

$$= \frac{5}{15} = \frac{1}{3}$$

இதுவே முழுமைத் தொகுதியில் நாம் குறிப்பிட்ட பண்பிற்கான நிகழ்திறனாக அமைகிறது. இம் முறைப்படி பார்ப்பின், முழுமைத் தொகுதிக்கான பண்பினை மாதிரியின் நிகழ்வெண் பரவலிலிருந்து திட்டமிடுவது நம்பிக்கை தருவதாய் அமைந்துள்ளதை உணர்கிறோம்.

மாதிரியிலிருந்து மதிப்பிடப்படும் முழுமைத் தொகுதி அளவை ஒரு குறிப்பிட்ட மதிப்பைக் குறிக்குமானால், அதைப்

புள்ளி மதிப்பீடு (point estimation) என்றும், முழுமைத் தொகுதி அளவை எந்த எல்லைக்குள் அமையும் என்பதை அம் முடிவினுக்குரிய நிகழ்திறனோடு இரு எல்லை எண்களைக் குறிப்பிடும்போது இடைவெளி மதிப்பீடு (interval estimation) என்றும் கொள்கிறோம். ஒரு சிறப்பான மதிப்பீடு பிறழ்ச்சியற்ற மதிப்பீடாகவும் (unbiased estimate), முரண்படா மதிப்பீடாகவும் (constant estimate), திறமுடைய மதிப்பீடாகவும் (efficient estimate), போதுமான மதிப்பீடாகவும் (sufficient estimate) அமையப்பெற வேண்டும்.

'θ_1', என்று கொடுக்கப்பட்ட மாதிரி அளவையின் கணித எதிர்பார்ப்பு (Mathematical expectation) அதற்கிசைந்த முழுமைத் தொகுதியின் அளவையான 'θ' ஆனுல், θ_1 என்பது θ-வின் பிறழ்ச்சியற்ற மதிப்பீடாகும். மாதிரியிலிருந்து கிடைக்கும் கூட்டுச் சராசரி ($\bar{x}$), அதன் முழுமைத் தொகுதிக்கான கூட்டுச் சராசரியின் (μ) ஒரு பிறழ்ச்சியற்ற மதிப்பீடாகும். ஆனுல்,

$$S^2 = \frac{\Sigma (x_i - \bar{x})^2}{n}$$

என்ற வாய்பாட்டினுல் கணிக்கப்படும் மாதிரியின் மாறுபாட்டளவை (S^2) முழுமைத் தொகுதிக்கான மாறுபாட்டளவை (σ^2)-க்குப் பிறழ்ச்சியற்ற மதிப்பீடாகாது. இதன் மாதிரியின் பரவலுக்கான (sampling distribution) கூட்டுச் சராசரி σ^2-க்குக் குறைவாக இருக்கும். S^2-பிறழ்ச்சியில்லாத மதிப்பீடாக அமையவேண்டுமானுல் $\Sigma (x_i - \bar{x})^2$ ஐ 'n'-க்குப் பதிலாக $(n-1)$ ஆல் வகுக்க வேண்டும். σ^2-ன் பிறழ்ச்சியில்லாத மதிப்பீடு $\Sigma (x_i - \bar{x})^2 / n-1$ என்பதை t-பரவலில் காண உள்ளோம்.

மாதிரியின் எண்ணிக்கையை (sample size) அதிகப்படுத்தும் போது, 'θ_1'-ன் மதிப்பு நிகழ்திறப்படி θ-க்கு நெருக்கமடைந்தால், மாதிரி அளவையான 'θ_1' என்பதை முழுமைத் தொகுதி அளவையான 'θ'-வுக்கு உறுதியான மதிப்பீடு(consistants estimate) என்கிறோம். மேற்சொன்ன விதியில் 'n' நிலையென்பதும் இவ் விதியில் 'n' வரம்பின்றி அதிகரிக்கப்படுவதும் உணரத்தக்கது.

ஒரு முழுமைத் தொகுதியின் இரு மாதிரி அளவைகளில் எந்த மதிப்பீட்டிற்கு மாறுபாட்டளவை (S^2) குறைவாக உள்ளதோ அதனையே திறமுடை மதிப்பீடு (efficient estimate) என்கிறோம். மிகக் குறைந்த மாறுபாட்டளவை உள்ள மதிப்பீடே மிகச் சிறந்த திறமுடையதாகும். சில குறிப்பிட்ட மாதிரிப் பரவல்களின் பண்புகளை ஆராயும்போது, அவற்றின் மாறுபாட்டளவைகள், திட்ட விலக்கங்கள் குறித்துப் பெரிதும் அக்கறை காட்ட

உள்ளோம். இவ் வளவைகள் திறனுக்கும், நம்பகத் தன்மைக்கும் அளவுகோலாக அமைவதால் புள்ளியியல் அனுமானத்தில் மிக்க முக்கியமுடைத்து.

மாதிரியினின்று பெறப்பட்ட அளவு சம்பந்தப்பட்ட முழுமைத் தொகுதிபற்றி எல்லாச் செய்திகளையும் தெரிவிப்பதாய் அமையின், அம் மதிப்பீடு போதுமான மதிப்பீடு (sufficient estimate) எனக் கொள்ளப்படும்.

இடைவெளி மதிப்பீடு : நம்பிக்கை வரம்புகள் (interval estimation confidence limits)

ஒரு முழுமைத் தொகுதியின் அளவைக்கு மாதிரிகள் அனைத்தும் ஒரே மதிப்பைக் கொடுக்கும் தன்மையாய் அமையாது. இவற்றில் எண்ணற்றவை தவறானவையாகவும், ஒரு மதிப்பீடு மட்டும் சரியானதாகவும் இருக்கும். மாதிரிகள் அனைத்தும் ஒரே மதிப்பைத் தராவிடத்து, நம்மால் மதிப்பிடப்படும் மதிப்பீட்டிற்கான தன்மையும் வேறுபடும். நம்மால் மதிப்பிடப்பட்ட மதிப்பீட்டிற்கான நம்பக எல்லைகளைத் (confidence intervals) தந்து, அம் மதிப்பீடு இவ் வெல்லைகளுக்குள் அமையும் நிகழ்திறத்தையும் குறிப்பிடுவது இடைவெளி மதிப்பீடு (Interval estimation) எனப்படும்.

கூட்டுச் சராசரியின் $(\bar{x})$ மாதிரிப் பரவல் μ என்ற கூட்டுச் சராசரியையும், $\sigma/\sqrt{n}$ என்ற திட்டப் பிழையையும் கொண்டு இயல்நிலைப் பரவலை ஒத்து அமையும். இயல்நிலைப் பரவலின் நியதியைக் கொண்டு $\bar{x}$ என்ற மாதிரி அளவை, μ என்ற அளவிலிருந்து எவ்வளவு மாறுபட்டிருக்கும் என்பதை நிகழ்திறப்படி குறிக்கிறோம். 100 முறைகளில் 95 முறை $(\mu-\bar{x})$ என்ற அளவு $-1{\cdot}96\,\sigma/\sqrt{n} < (\mu-\bar{x}) < 1{\cdot}96\sigma/\sqrt{n}$ என்ற பிரிவில் அமையும் எனக் கொள்கிறோம்.

$$-1{\cdot}96\sigma/\sqrt{n} < (\mu-\bar{x}) < 1{\cdot}96\,\sigma/\sqrt{n}$$
$$\bar{x}-1{\cdot}96\;\;\sigma/\sqrt{n} < \mu < \bar{x}+1{\cdot}96\,\sigma/\sqrt{n}$$

$1{\cdot}96$ என்பது $2{\cdot}58$ என மாற்றப்படும்போது இடைவெளியின் அளவு அதிகமாவதோடு நமது மதிப்பு இவ் விடைவெளியில் அமைவதற்கான எண்ணிக்கையும் 100-க்கு 99 என உயர்த்தப் படுகிறது. இதனையே நிகழ்திற நியதிப்படி குறிப்பிட்டால், மதிப்பீடானது $\bar{x}\;\pm\;1{\cdot}96\,\sigma\sqrt{n}$ என்ற இடைவெளியில் அமைவதற்கான நிகழ்திறன் 0.95 $\pm\;2{\cdot}58\,\sigma/\sqrt{n}$ என்ற

இடைவெளியில் அமைவதற்கான நிகழ்திறன் 0·99 ஆகவும் இருக்கும் எனச் சொல்கிறோம்.

கூட்டுச் சராசரிக்கு மாறாக, முழுமைத் தொகுதியின் எந்த அளவைக்கும் (திட்ட விலக்கம், மாறுபாட்டளவை, உடன் தொடர்பு கெழு......) அந்தந்த அளவைகளின் திட்டப் பிழைகளின் துணைகொண்டு நம்பிக்கை எல்லைகளை வரையறுக்க முடியும்.

கருதுகோள் சோதனைகள் (Testing of Hypothesis)

ஒரு முக்கோணத்தின் மூன்று கோணங்களின் (பாகைகளின்) கூட்டுத் தொகை 180° என ஒரு கருதுகோளை (hypothesis) ஏற்படுத்தினால், இதற்கான வரைபடம் வரைந்து, இக் கருதுகோள் உண்மையா, பொய்யா என்பதைச் சோதனை செய்யலாம். இக் கருதுகோளுக்கான கணக்கியல் நிரூபணம் அமைந்துள்ளது. நிரூபணம் முடிந்தபின் கருதுகோளின் தன்மைபற்றி நமக்கு உறுதி ஏற்படுகிறது. புள்ளியியல் துறையில் இதுபோன்ற கருதுகோள்களைக் காண்பது மிக அரிதாகும். முழுமைத் தொகுதியின் ஓர் அளவையைப்பற்றிச் சொல்லப்படும் கருத்தே புள்ளியியல் கருதுகோள் (statistical hypothesis) எனப்படும். முழுமைத் தொகுதி முழுமையும் சோதனை செய்த பின்பே இக் கருதுகோள் நம்பும் தன்மைத்தா இல்லையா என்ற முடிவுக்கு வரவேண்டும். எல்லாவிடத்தும் முழுமைத் தொகுதியைச் சோதனை செய்வது சாத்தியமன்று. முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட ஒரு மாதிரியைக்கொண்டு முழுமைத் தொகுதிக்கான கருதுகோள் உண்மையா இல்லையா என்பதைச் சோதிக்க வேண்டியவர்களாக உள்ளோம். மாதிரியின் அளவைகளைக் கொண்டு கருதுகோள் சோதனை செய்து, கருதுகோள் நம்பும் தன்மைத்தா இல்லையா என்ற முடிவிற்கு வருவது கருதுகோள் சோதனை முறையாகும். இம்மாதிரி எடுக்கப்பட்ட முடிவுகளில் தவறு இராது என உறுதியாகக் கூற முடியாது. இம் முடிவுகளில் இருவகையான தவறுகள் ஏற்படுவதற்கு இடமிருக்கிறது. கருதுகோள் உண்மையாக இருக்கும்போது மறுத்தும் (rejecting the hypothesis when it is true), கருதுகோள் தவறாக இருக்கும் போது ஏற்றும் (accepting the hypothesis when it is false) இரு வழிகளில் தவறிழைக்க நேரிடும். ஒரு சோதனையில் தவறான முடிவுகளை ஏற்படுத்துவதற்கான நிகழ்திறங்களைக் குறிப்பிடுவதே புள்ளியியல் கருதுகோள்களின் சிறப்பம்சமாகும். இத் தன்மை எந்தத் தவற்றினை இழைப்பது அபாயத்திற்கிடமானதாக உள்ளதோ அதைக் குறைத்து அமைப்பதற்கான முறைகள் அமைந்திருப்பது மற்றொரு சிறப்பம்சமாகக் கொள்ளலாம்.

புள்ளியியலில் அமைக்கப்படும் கருதுகோள்கள் யாவும் மறுக்கத்தகு கருதுகோள்களாக (null hypothesis) அமைக்கப்படும். 'N' உறுப்புகள் கொண்ட முழுமைத் தொகுதியிலிருந்து n உறுப்புகள் கொண்ட மாதிரி எடுப்பதாகக் கொள்வோம். முழுமைத் தொகுதிக்கான கூட்டுச் சராசரியை μ எனவும் மாதிரியின் கூட்டுச் சராசரியை $\bar{x}$ எனவும் குறிப்பிடுவோம். n உறுப்புகள்கொண்ட மாதிரியை, முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட ஓர் இயைபிலா மாதிரியாகக் கொண்டு, ஒரு கருதுகோள் அமைப்போம். $\bar{x}$ என்ற கூட்டுச் சராசரி முழுமைத் தொகுதிக்கான கூட்டுச் சராசரி μ-வைக் குறிக்கும் என்பதே கருதுகோளாகக் கொள்ளப்படும். இதனையே புள்ளியியல் கருதுகோளாக அமைக்கும்போது இவ்விரு சராசரிகளுக்குமான வித்தியாசம் $(\mu - \bar{x})$ குறிப்பிடத்தக்கதாய் இல்லை எனக் கொள்கிறோம். (There is no significant difference between the two means.) இக் கருதுகோள் உண்மையாக இருக்கும்போது, இதில் ஏற்பட்டுள்ள வித்தியாசம் (difference between two means) கூறெடுத்தலின் ஏற்றத்தாழ்வுகளுக்கானது(sampling fluctuations) எனக் கொள்கிறோம் இக் கருதுகோள் உண்மையாக இருக்கும் போது நிகழ்வெண் பரவலில் அமைந்த எண்களைக் கொண்டு, கருதுகோள் சோதனையில் கிடைக்கும் எண்களோடு ஒப்பிட்டு முடிவு செய்கிறோம். ஏற்படுத்தப்பட்ட கருதுகோள் மறுக்கப் படும்படி அமைவதால், இவை யாவும் மறுக்கத்தகு கருதுகோள் (null hypothesis) எனப்படும். இக் கருதுகோள் மறுக்கப்படும் போது, இவ்விரு கூட்டுச் சராசரிகட்கும் உள்ள வேறுபாடு குறிப்பிடத்தக்கதாய் உள்ளது எனக் கொள்கிறோம்.

புள்ளியியல் கருதுகோள்கள் யாவும் அடியிற்காணும் ஏதாவது ஒரு பண்பைப் பற்றியதாய் அமையலாம்:

(1) ஒரு மாறி அல்லது பண்பிற்கான அளவிற்குரிய (magnitude) அனுமானமாகவோ

(2) இரண்டளவைகளுக்கான வித்தியாசத்தைக் (difference) குறிக்கும் அனுமானமாகவோ

(3) இரு மாறிகள் அல்லது பண்புகளுக்கான உடன்தொடர்பைக் குறிக்கும் அனுமானமாகவோ

(4) மதிப்பீடு செய்யப்பட்ட யுத்திகளின் தன்மைக்கான அல்லது அவ் வுத்தியின் நிகழ்வெண் பரவலுக்கான அனுமானமாகவோ, அல்லது

(5) மாறி அல்லது பண்பிற்கான போக்கைக் குறிக்கும் அனுமானமாகவோ இருக்கலாம்.

அளவைக் குறித்த அனுமானத்திற்கு இவ்வாண்டு நம் நாட்டில் விளையும் நெல்லின் மொத்த அளவையையோ அல்லது வேலையில்லார்தம் எண்ணிக்கையையோ குறிப்பிடலாம். சிறந்த அனுமானம் எல்லாப் பண்புகளையும் பெற்ற மதிப்பீடாக அமைவதால், இவ் வனுமானமும் மதிப்பீட்டு முறையை ஒத்ததாகும். இரண்டளவைகளுக்கான வித்தியாச அனுமானத்திற்கு, இரண்டு மாநிலத்தில் வயது வந்தோர்தம் உயரத்தின் வித்தியாசத்தையும் அல்லது ஒத்த வயதுடைய இரு பாலார்க்கு (both sex) உள்ள அறிவுக் குறியீட்டெண்ணின் வித்தியாசத்தையும் சோதனை செய்வதைக் குறிப்பிடலாம். இரு மாறிகளின் உடன் தொடர்பிற்கான அனுமானத்திற்கு உயரத்துக்கும் எடைக்கும் அல்லது நீரின் அளவிற்கும் விளைபலனுக்கும் உள்ள உடன் தொடர்பின் அனுமானத்தைக் கூறலாம். ஒரு யுத்தியின் திறமை அல்லது திறமையற்ற தன்மைக்கான அனுமானமாகத் தற்சமயம் கொடுக்கப்படும் உரத்தின் அளவை மாற்றுவதாலோ, வேறு உரம் இடுவதாலோ உண்டாகும் விளைச்சலின் தன்மையைக் குறிப்பிடலாம். மாறியின் போக்கின் தன்மைக்கான காட்டாக, நமது நாட்டின் உணவு உற்பத்தி கடந்த சில ஆண்டுகளாக எப்படி உள்ளது அல்லது நமது தினசரி உணவு உட்கொள் அளவின் தன்மை எப்படி மாறுபட்டுள்ளது என்பதைக் கூறலாம்.

மேற்சொன்ன கருத்துகள் புள்ளியியல் முறைப்படி சோதனை செய்யுமுன் ஒன்று அல்லது அதற்கு மேற்பட்ட புள்ளியியல் கருதுகோட்களாக மாற்றியமைக்கப்பட வேண்டும். புள்ளியியல் கருதுகோள் என்பதை, மாதிரியைக் கொண்டு மதிப்பிட்ட ஓர் அளவைக்கு, நாம் கொடுக்கும் நம்பகத்திற்கான நிகழ்திறனை நிர்ணயிக்கும் ஒரு நிகழ்வெண் சார்பிற்கான (frequency function) தன்மையைக் குறிக்கும் கருத்தாகக் கொள்ளலாம்.

புள்ளியியல் கருதுகோள் உண்மையாக இருக்கும்போது, சோதனை அளவை (test statistic), ஒரு குறிப்பிட்ட அளவினைப் பெறுவதற்கான நிகழ்திறனைக் கொடுக்கும் நிகழ்வெண் பரவலைப் பெற்றிருக்கும். இவ் வளவைகள் யாவும் மாதிரியிலிருந்து கணக்கிடப்படும் கூட்டுச் சராசரி ($\bar{x}$), மாறுபாட்டளவை (S^2), திட்டவிலக்கம் (S), உடன்தொடர்பு கெழு (r) மற்றும் இது போன்ற அளவைகளால் பெறப்படுகிறது. சில சோதனை அளவைகள் வேறுபட்ட கருதுகோள்களை அனுமானிக்க உபயோகப்படுத்தப் பட்டாலும், சில அளவைகளின் உபயோகம் மிகவும் குறைந்ததாக அமைந்துள்ளது. ஒவ்வோர் அளவையும் சில குறிப்பிட்ட பண்பைப் பெற்றதாயும், குறிப்பிட்ட முறையில் அமைக்கப்பட்ட

தாயும் இருக்குமெனக் கொண்டு முடிவுகள் எடுக்கப்படுகின்றன. நமக்குத் தேவையான சில முக்கியமான சோதனை அளவைகளை அடியிற்காணும் ஐந்து பரவல்களைக் கொண்டு திட்டமிட உள்ளோம்:

1. ஈருறுப்புப் பரவல் (Binomial distribution)
2. இயல்நிலைப் பரவல் (Normal distribution)
3. t-பரவல் (t-distribution)
4. கை-வர்க்கப் பரவல் (Chi-square distribution)
5. F-பரவல் (F-distribution).

கருதுகோளை ஏற்பதா மறுப்பதா என நம்மால் எடுக்கப்படும் முடிவு மாதிரிகளின் அளவைகளைச் சார்ந்ததாயும் இம்முடிவை எடுக்கும்போது, நாம் தவறான முடிவுகளை மேற்கொள்வதற்கான நிகழ்திறனைக் குறிப்பிடுவதாயும் அமைகிறது. முழுமைத் தொகுதிக்குரிய ஓர் அளவைக்கான கருதுகோளை அமைப்பதாகக் கொள்வோம். ஒரு மாதிரியினை எடுத்து, அதன் உறுப்புகளின் தன்மையை அறிந்து, இவை நமது முழுமைத் தொகுதிக்கான அளவையை ஒத்திருக்கிறதா இல்லையா என்ற முடிவிற்கு வர உதவியாக இருக்கும் சோதனை அளவைக் கணக்கிடுகிறோம். மாதிரியின் அளவை, முழுமைத் தொகுதியின் அளவைக்கு மிக நெருக்கமாக இருப்பின் கருதுகோள் ஏற்கப்பட்டும், நெருக்கமாக அமையாதவிடத்துக் கருதுகோள் மறுக்கப்பட்டும் முடிவு செய்யப்படும்.

சிறப்பு வரம்பு (Level of significance)

கருதுகோள் சோதனை செய்யும் முறையை கூட்டுச் சராசரிக்கான நிகழ்வெண் பரவலால் விளக்கலாம். முழுமைத் தொகுதியின் கூட்டுச் சராசரியான $\mu = 50$ என்பது நமது கருதுகோளாக இருக்கட்டும். n உறுப்புகள் கொண்ட மாதிரி ஒன்றிற்கான கூட்டுச் சராசரியைக் $(\bar{x})$ கணக்கிடுவோம். நமது கருதுகோள் உண்மையாக இருக்குமிடத்து, மாதிரியின் கூட்டுச் சராசரி 50-லிருந்து மிக விலகி அமைவது மிக்க அரிதென்பதைக் கூட்டுச் சராசரிக்கான நிகழ்வெண் பரவலால் உணர்கிறோம். மாதிரியின் கூட்டுச் சராசரி 50-க்கு நெருக்கமாக இருப்பின் கருதுகோளை ஏற்றும், மிக விலகி இருப்பின் கருதுகோளை மறுத்தும் முடிவுகளை மேற்கொள்கிறோம். 10 என்ற திட்ட விலக்கம் கொண்ட முழுமைத் தொகுதியிலிருந்து 25 உறுப்புகள் கொண்ட மாதிரி எடுத்தால், கூட்டுச் சராசரிக்கான மாதிரியின்

பரவல் $\sigma/\sqrt{n} = 10/\sqrt{25} = \frac{10}{5} = 2$ என்ற திட்டப் பிழையுடன் இயல்நிலைப் பரவலை ஒத்து அமையும். முழுமைத் தொகுதிக்கான கூட்டுச் சராசரியும், மாதிரியின் கூட்டுச் சராசரியும் 50 ஆக அமையும்.

மாதிரியின் கூட்டுச் சராசரிக்கான எந்த மதிப்புகளுக்குக் கருதுகோள் மறுக்கப்படும் என்பதைப் பார்ப்போம்,

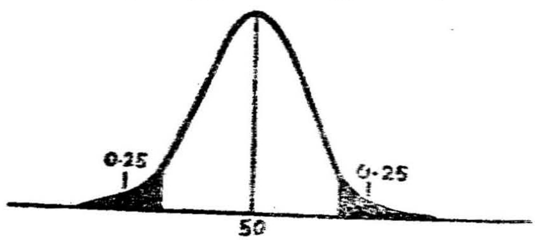
$$\bar{x} = 50 \pm 1{\cdot}96\, \sigma/\sqrt{n}$$

$$= 50 \pm 1{\cdot}96 \times 2$$

$$= 50 \pm 3{\cdot}92$$

$$= 46{\cdot}08,\ 53{\cdot}92$$

என்ற எல்லைகளைக் கடந்து அமையும் அளவைகளைக் கொண்ட $\bar{x}$' கிடைக்குமிடத்துக் கருதுகோள் மறுக்கப்படும். $\bar{x}$, 46·08-க்குக் குறைவாகவும், 53·92-க்கு அதிகமாகவும் மதிப்பைப் பெறும் போது $\mu = 50$ என்ற கருதுகோள் மறுக்கப்படும்.



படம் 10·1

இயல்நிலைப் பரவலின் தன்மைப்படி, கூட்டுச் சராசரிக்கு இருபுறமும் 1·96 திட்டவிலக்க தூரம் 95 சத அளவுகளை உட்கொள்ளும் என்பதை உணர்வோம். இதனால் $\mu = 50$ என்ற கருதுகோள் 100-க்கு 5 முறைகள் ($\alpha = {\cdot}05$) உண்மையாக இருக்கும்போது மறுக்கப்படும் தன்மைத்தாய் உள்ள நிலையை உணர்கிறோம். கூட்டுச் சராசரி ஐம்பதும், திட்டவிலக்கம் பத்தும் உள்ள முழுமைத் தொகுதியிலிருந்து 25 உறுப்புகள் கொண்ட பல மாதிரிகள் எடுப்பதாகக் கொள்வோம். நூறு மாதிரிகளில் ஐந்து மாதிரிகளின் அளவு 46·08-க்குக் குறைவாகவும், 53·92-க்கு அதிகமாகவும் அமைவதால், கருதுகோள் உண்மையாக இருக்கும்போது எடுக்கப்படும் நூறு முடிவுகளில் ஐந்து முடிவுகள் தவறாக இருக்கும் என்ற அபாயநிலையை

உணர்ந்து செயல்படுகிறோம், 95 முறைகள் $\bar{x}$-ன் மதிப்பு (46·08, 53·92) என்ற எல்லைக்குள் அமைவதால், இவ்விடங்களில் நமது முடிவுகள் தவறில்லாமல் இருக்குமென்ற நம்பிக்கை இருக்கிறது. செயல்முறையில் ஒரு தனித்த மாதிரியைக் கொண்டு முடிவுகள் எடுக்கப்படுவதால் இந்த அபாயநிலை (100-க்கு 5) ஒவ்வொரு முடிவிற்கும் பொருந்தும்.

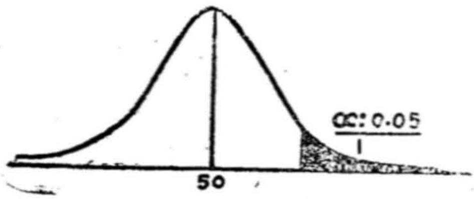
கருதுகோளை மறுப்பதற்கான $\bar{x}$-ன் மதிப்புகளே மறுப்பு மண்டிலம் (rejection region) அல்லது மாறுநிலை மண்டிலம் (critical region) எனப்படும். இக் கணக்கில் $\bar{x}$-ன் மறுப்பு மண்டிலம் $\mu = 50$ ஆக உள்ள கருதுகோளுக்கு $\bar{x} < 46{\cdot}08$ மற்றும் $x > 53{\cdot}92$ என அமைகிறது. கருதுகோள் ($\mu = 50$) என்பது உண்மையாக இருக்கும்போது ஒரு மாதிரிக்கான அளவை இம் மறுப்பு மண்டிலத்தில் அமைவதற்கான நிகழ்திறம் 0·05. 100 மாதிரிகளில் 5 மாதிரிகளின் மதிப்புத்தான் இம் மண்டிலத்தில் அமையும் அளவைகளைப் பெறும். இதனையே சோதனையின் சிறப்பு வரம்பு (level of significances) என்கிறோம். இச் சிறப்பு வரம்பு 0 05 இருக்கும் தன்மைக்கான நிலையை மறுக்கப்படும் புள்ளியியல் கருதுகோளுக்கான மதிப்புக் கவனிக்கப்படத்தக்க சிறப்புடை (statistically significant) நிலையில் உள்ளதென்றும் இச் சிறப்பு வரம்பு 0·01 என இருக்கும் நிலையில் மறுக்கப்படும் புள்ளியியல் கருதுகோளுக்கான மதிப்பு மிகவும் சிறப்புடையதாய் (highly significant) குறிப்பிடத்தக்க வித்தியாசத்தைப் பெற்றிருக்கிறது எனவும் கொள்கிறோம்.

ஒரு வால்புறச் சோதனை (One tail test)

சில சமயங்களில் கூட்டுச் சராசரி, ஒரு திசையில்மட்டும் ஒரு குறிப்பிட்ட அளவிற்குமேல் அமைந்தால் கருதுகோள் மறுக்கப்படுவது சிறந்ததாய் அமைகிறது. கருதுகோள் $\mu = 50$ என உள்ளபோது $\sigma = 10$, $n = 25$, $\lambda = 05$ என்ற அமைப்பால் ஒரு வால்புறச் சோதனைக்கு $\cdot 50 + 1{\cdot}645\, \sigma/\sqrt{n} = 50 + 1{\cdot}645 \times 2$ $53{\cdot}39$ என்ற எல்லைக்குமேல் அமைந்த அளவைகள் மறுப்பு மண்டிலத்தில் அமையுமெனக் கொள்கிறோம்.

ஒருபுற அமைப்பு மண்டிலத்தை உபயோகப்படுத்தும்போது 50-க்கு மேற்பட்ட கூட்டுச் சராசரி அளவைகளைக் கொண்ட மாறுபட்ட முழுமைத் தொகுதிகளையே கவனிக்கிறோம். கூட்டுச் சராசரியின் அளவு 50-க்குக் குறைவாக அமைந்தால் கருதுகோள் ஏற்கப்படுவதையும், அதன் அளவு ஓர் எல்லைக்கு

மேல் அமையும்போது கருதுகோள் மறுக்கப்படுவதையும் இம் முறை தெளிவுபடுத்துகிறது.



படம் 10·2

ஒரு வால்புறச் சோதனைக்கான சில எடுத்துக்காட்டுகளைக் குறிப்பிடுவோம். குடிநீரில் உள்ள நோய்க் கிருமிகளுக்கான சோதனையில், குடிநீர் குடிப்பதற்காகாது என்ற நிலையை நோய்க்கிருமிகளின் எண்ணிக்கை ஒரு குறிப்பிட்ட எண்ணைத் தாண்டியவுடன் முடிவு செய்யப்படலாம். புதுமாதிரியான உரத்தின் தரம்காணச் செய்யப்படும் சோதனையில், பழைய உரத்தினால் கிடைத்த விளைச்சலைக் காட்டிலும் ஒரு குறிப்பிட்ட அளவு அதிகம் கொடுப்பதைக்கொண்டே முடிவு செய்கிறோம். எனவே, எந்த அளவுக்கு அதிகமான அல்லது குறைவான மதிப்பைமட்டும் எல்லையாகக் கொள்வது சிறப்பாக அமையுமோ அவ்விடங்களில் ஒரு வால்புறச் சோதனைகளை மேற்கொள்வது நன்மை பயப்பதாய் அமைகிறது.

ஒரு கருதுகோட் சோதனைக்கான தன்மையை அடியிற் காணும் முறைப்படி செயல்படுத்துவது முறையாகும்.

1. கருதுகோளை அமைத்தல்

2. கருதுகோளை ஏற்க அல்லது மறுக்க உதவும் சிறப்பு வரம்புகளை வரையறுத்துக் கொள்ளுதல்

3. மாதிரியின் அளவுகளைக் கொண்டு சோதனை அளவையைக் (test statistics) கணக்கிடுதல்

4. சிறப்பு வரம்புகளுக்கான (level of significance) நிகழ்திற மதிப்புகளுடன் சோதனை அளவைகளை ஒப்பிட்டுக் கருதுகோளை ஏற்றும் மறுத்தும் முடிவுசெய்தல்.

பயிற்சி :

1. புள்ளி மதிப்பீட்டிற்கும் இடைவெளி மதிப்பீட்டிற்குமான வேற்றுமை யாது?

2. கொடுக்கப்பட்ட அளவுகளுக்கான மூன்று உறுப்புகள் கொண்ட மாதிரிகள் முழுதும் அமைத்து, அதற்கான கூட்டுச் சராசரி மதிப்பீடுகள் ($\bar{x}$) முழுமைத் தொகுதிக்கான மதிப்பீட்டுடன் (μ) ஒத்து அமைகிறதா என்பதைச் சோதனை செய். இதற்கான நம்பக எல்லைகளை அமைக்கும் முறையை விளக்கு.

49	54	43
57	24	55
16	95	55
78	64	56
09	47	27

3. கருதுகோள் சோதனையின் திட்ட முறையைத் தக்க சான்றுகளுடன் விளக்கு.

4. மறுக்கத்தகு கருதுகோள் அமைக்கும் தன்மையையும், இதனை மறுப்பதாலும் ஏற்பதாலும் கிடைக்கக்கூடிய நிகழ்திறன்களின் தன்மையையும் தக்க எடுத்துக்காட்டுகளுடன் விளக்கு.

5. ஒரு வால்புறச் சோதனையின் தன்மை யாது? இச்சோதனை அமைக்கப்பட வேண்டிய இடங்களின் தன்மைபற்றித் தக்க சான்றுகளுடன் விவரி.

11. சிறப்புக்காணும் சோதனைகள்

(Test of Significance)

(ஈருறுப்பு இயல்நிலைப் பரவல்களைத் தழுவியது)

சிறப்புக்காணும் சோதனைகள் : ஈருறுப்புப் பரவல்

நிகழ்ச்சிக்கு நிகழ்ச்சி நிகழ்திறன் மாறுபடும் தன்மையில்லாத சார்பற்ற 'n' நிகழ்ச்சிகளில், 'x' வெற்றிகளுக்கான நிகழ்திறனைக் குறிக்க ஈருறுப்புப் பரவலை உபயோகப்படுத்தலாம் என்பதைக் கண்டோம். இப் பரவலின் துணைகொண்டு பண்புகளின் அடிப் படையில் அனுமானிக்கப்படும் மாறிகளின் அளவைகளுக்கான சில சிறப்புக் காணும் சோதனைகளை அமைப்போம்.

நமது கருதுகோள் ஒரு நிகழ்ச்சி நிகழ்வதற்கான நிகழ்திறன் 'p' என்பதாக இருக்கட்டும். நமது கருதுகோள் உண்மையாக இருக்கும்போது, வெற்றிகளின் எண்ணிக்கைக்கான மாதிரியின் பரவல் (sampling distribution), np கூட்டுச் சராசரியையும் npq மாறுபாட்டளவையையும் கொண்டிருக்கும். n அதிகமாக்கப்படும் போது $z = \dfrac{x-np}{\sqrt{npq}}$ என்ற மாறி இயல்நிலைப் பரவலை ஒத்து அமையும் என்பதைக் குறிப்பிட்டோம். z-ன் மதிப்பை ஒரு சோதனை அளவையாகக்கொண்டு, நிகழ்ச்சி நிகழ்வதற்கான நிகழ்திறன் 'p' என்ற கருதுகோளைச் சோதனை செய்கிறோம்.

கருதுகோள் உண்மையாக இருக்கும்போது, z-ன் இருபுறமும் 1·96 திட்ட விலக்க தூரம் 95 சத அளவுகளும், z-ன் இருபுறமும் 2·58 திட்டவிலக்க தூரம் 99 சத அளவுகளும் அமையும் என்பதை அறிவோம்.

$$p(-1{\cdot}96 \le z \le 1{\cdot}96) = p(|z| \le 1{\cdot}96) = 0{\cdot}95$$

$$p(-2{\cdot}58 \le z \le 2{\cdot}58) = p(|z| \le 2{\cdot}58) = 0{\cdot}99$$

நிகழ்ச்சிகளின் மதிப்புகள் ($-1{\cdot}96$, $+1{\cdot}96$) இடைவெளிக்கப்பால் அமையும்போது, அதன் நிகழ்திறன் $\cdot 05$-க்குக் குறைவாகவும், மதிப்புகள் ($-2{\cdot}58$, $+2{\cdot}58$) இடைவெளிக்கப்பால் அமையும்போது அதன் நிகழ்திறன் $0{\cdot}01$-க்குக் குறைவாகவும் மதிப்புகளைப் பெறும். நூற்றுக்கு ஐந்து, நூற்றுக்கு ஒரு நிகழ்ச்சி தான் இவ் விடைவெளிகளுக்கு வெளியில் அமையும் நிகழ்திறத்தைப் பெறுகிறது. இந்த 5%, 1% சிறப்புக்காணும் எல்லைகளைக் கொண்டே பெரும்பாலான சோதனைகளுக்கான முடிவுகள் மேற்கொள்ளப்படுகின்றன. வேறு நிகழ்திற எல்லைகளை அமைத்துக் கொள்வதில் தவறேதுமில்லை. இதிலும் 5% நிகழ்திற எல்லையைக் கொண்டு கருதுகோளை ஏற்றோ, மறுத்தோ முடிவு செய்வதே நடைமுறையில் பெரும்பாலும் காணப்படுகிறது.

z-ன் தனி மதிப்பை (absolute value) அதாவது $|z|$ ஐக் கணக்கிட்டு, அம் மதிப்பு $1{\cdot}96$ அல்லது அதற்குக் குறைவான மதிப்பினைப் பெற்றால் கருதுகோளை ஏற்றும், $1{\cdot}96$-க்கு அதிகமான மதிப்பினைப் பெற்றால் கருதுகோளை மறுத்தும் சிறப்புக்காணும் எல்லை 5% ஆக உள்ளவிடத்து முடிவுசெய்யப்படுகிறது. $\frac{(x-np)}{npq} = |z|$ ஐ $1{\cdot}96$, $2{\cdot}58$ உடன் ஒப்பிடுவதற்கு மாறாக $(x-np)$ ஐ $1{\cdot}96$, $2{\cdot}58$ திட்டப்பிழை மதிப்புடன் ஒப்பிட்டு முடிவு செய்யப்படுவதும், சோதனை அளவை $|z|$ ஐ, $1{\cdot}96$, $2{\cdot}58$ உடன் ஒப்பிடுவதற்கு மாறாக 2, 3 உடன் ஒப்பிட்டு முடிவு செய்யப்படுவதும் ஒரே முடிவுகளைத் தரும். பெரிய அளவு மாதிரிகளுக்கான சோதனையில் (large sample tests) 2, 3 உடன் ஒப்பிட்டு முடிவுசெய்வதே நடைமுறையில் உள்ளது.

காட்டு :

இரு இனப் பட்டாணிகளைக் கலக்கச் செய்ததால் (cross) கிடைத்த விதைகளில் 5321 மஞ்சளாகவும், 1804 பச்சையாகவும் அமைந்துள்ளன. இம் முழுமைத் தொகுதியில் மெண்டலியன் கருதுகோட்படி 25 சதப் பச்சை நிறப் பட்டாணி எதிர்பார்க்கப் பட்டால், இம் மாதிரியில் அமைந்துள்ள பச்சைப் பட்டாணிக்கான விகித சம வித்தியாசம் (deviation from 25%) மாதிரிளடுத்தலின் ஏற்றத் தாழ்வாகக் கொள்ளலாமா என்பதற்கான சோதனையைக் காண்போம்.

சோதனையில் கிடைத்த பச்சைநிறப் பட்டாணிகள் = 1804.

$$\left.\begin{array}{l}\text{மெண்டீலியன் முறைப்படி எதிர்பார்க்கப்} \\ \text{படும் பச்சைநிறப் பட்டாணிகள்}\end{array}\right\} = 7125 \times \frac{25}{100}$$

$$= 1781 \cdot 25$$

$$= 1781$$

$$\text{மாதிரியின் திட்டப்பிழை} = \sqrt{npq} = \sqrt{(7125)(\cdot 25)(\cdot 75)}$$

$$= 36 \cdot 6$$

$$\text{சோதனை அளவை} = z = \left| \frac{x - np}{\sqrt{npq}} \right|$$

$$= \left| \frac{1781 - 1804}{36 \cdot 6} \right| = 0 \cdot 65$$

$$< 1 \cdot 96$$

மெண்டீலியன் எதிர்பார்ப்பைவிட 23 விதைகள் அதிகமாக அமையப்பெற்றது. மாதிரியின் ஏற்றத் தாழ்வுக்கானதாகக் (sampling fluctuation) கொண்டு, கொடுக்கப்பட்ட மாதிரி மெண்டீலியன் கருதுகோளுக்கான முழுமைத் தொகுதியிலிருந்து வந்ததென முடிவு செய்கிறோம்.

இச் சோதனையைச் சற்று விரிவாகக் காண்போம். மெண்டீலியன் கருதுகோட்படி அமையும் முழுமைத் தொகுதியில் 25 சதம் பச்சைப் பட்டாணிகள் கிடைக்குமென நம்புகிறோம். 7125 பட்டாணிகள் கொண்ட ஒரு மாதிரியில் 1804 பச்சைநிறப் பட்டாணிகள் உள்ளதெனக் கணக்கிட்டுள்ளோம். நாம் எடுத்துள்ள மாதிரி மெண்டீலியன் கருதுகோளுக்கு ஒத்து அமைகிறதா என்ற சோதனை, நமது மாதிரியிலும் 25 சதப் பச்சை நிறப் பட்டாணிகள் உள்ளனவா என்பதைச் சோதிப்பதற்குச் சமமாகும். எடுக்கப்பட்ட மாதிரிகளில் சரியாக 25 சதப் பச்சைநிறப் பட்டாணிகள் அமைவது சாத்தியமன்று. 25 சதத்திற்குச் சற்று கூடியோ குறைந்தோ அமையப்பெறலாம். இம் மாதிரியில் 25 சத அளவிற்கு அதிகமாக 23 பட்டாணிகள் உள்ளன. நமது சோதனைக்கான கருத்து, இந்த வித்தியாசமான 23 விதைகள் மாதிரி எடுத்த முறையினால் ஏற்பட்ட ஏற்றத்தாழ்வுகளினால் (sampling fluctuations) ஏற்பட்டதா அல்லது குறிப்பிட்ட மெண்டீலியன் முழுமைத் தொகுதியிலிருந்து வந்தது எனக் கொள்ளும்படியாக அமையாமல், ஏற்பட்டுள்ள வித்தியாசம் (+23) குறிப்பிடத்தக்கபடி அமைந்துள்ளதா என்பதாகும். வித்தியாசம் குறிப்பிடத்தக்கபடி அமையவில்லை என்ற முடிவினை மேற்கொள்வோமாயின், மாதிரியின் அளவைக் கூட்டுவதாலோ அல்லது மாதிரிகளை அதிகப்

படுத்துவதாலோ வித்தியாசம் குறைந்து மெண்டலியன் முறையை ஒத்து அமையும் என நம்புகிறோம். மறுக்கப்படும் இடத்து இத் தன்மையை அடையாது என்ற கருத்திற்கான முடிவைச் சொல்கிறோம். 1·96 மற்றும் 2·56 உடன் ஒப்பிட்டு முடிவிற்கு வருவது, நம்மால் எடுக்கப்பட்ட முடிவினுக்குரிய நம்பக நிகழ்திற (·95, ·99) மதிப்பைக் காட்டுகிறது. 5% சிறப்பெல்லையில் (1·96 உடன் ஒப்பிடுவது) எடுக்கப்படும் முடிவுகளில் நூற்றுக்கு ஐந்து முறை நாம் தவறான முடிவிற்கு வருவதற்கான நிகழ்திறனும், 1% சிறப்பு எல்லையில் (2·56 உடன் ஒப்பிடுவது) எடுக்கப்படும் முடிவுகளில் நூற்றுக்கு ஒருமுறை தவறான முடிவிற்கு வருவதற்கான நிகழ்திறனும் அமைந்துள்ளது. தவறான முடிவிற்கு வருவதற்கான நிலையில் ஏன் முடிவு செய்ய வேண்டும் என எண்ணலாம். மற்றெந்த முறையில் எடுக்கும் முடிவில் தவறிராது? நமது தவறான முடிவிற்கான நிகழ்திறன் குறிப்பிடுவதில்தான் இதன் சிறப்பம்சம் அமைந்துள்ளது.

புள்ளியியல் கருதுகோள்கள் யாவும் இலையெனும் கருதுகோள் அல்லது மறுக்கத்தகு கருதுகோள்கள் (null hypothesis) ஆகும். எந்தச் சொற்களால் கருதுகோள் அமைக்கப்பட்டாலும் சோதனை செய்யப்படும் கருதுகோள் இலையெனும் அல்லது மறுக்கத்தகு கருதுகோளையே ஆகும். கொடுக்கப்பட்டுள்ள மாதிரிக்கான கருதுகோளைப் பார்ப்போம். மாதிரியில் உள்ள பச்சைநிறப் பட்டாணிகளின் விகித சமம் 25 சதத்திலிருந்து குறைந்தோ, கூடியோ இருப்பதை மாதிரி எடுத்தலின் ஏற்றத் தாழ்வுகளுக்கானதாகக் (sampling fluctuations) கொள்ளலாமா என்பதே கருதுகோட் சோதனையாக அமைந்துள்ளது. இதனையே 'மாதிரி 25 சதம் பச்சைநிறப் பட்டாணி தரும் மெண்டலியன் முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதெனவோ அல்லது மாதிரியில் உள்ள பச்சைநிறப் பட்டாணிகளுக்கும் மெண்டலியன் முழுமைத் தொகுதியில் எதிர்பார்க்கப்படும் பச்சைநிறப் பட்டாணிகளுக்கும் உள்ள வித்தியாசம் குறிப்பிடத்தக்கபடி இல்லை' எனவோ அமைத்துக் கொள்ளலாம்.

இச் சோதனைக்கான புள்ளியியல் இலையெனும் கருதுகோள் (statistical null hypothesis) 'மாதிரியில் உள்ள பச்சைநிறப் பட்டாணிகளின் விகித சமத்திற்கும், மெண்டலியன் முழுமைத் தொகுதியிலுள்ள பச்சைநிறப் பட்டாணிகளின் விகித சமத்திற்கும் உள்ள வித்தியாசம் குறிப்பிடும்படி இல்லை' என்பதாகும். கருதுகோள் எச் சொற்களால் அமைக்கப்பட்டாலும், சோதனை செய்ய மனத்தில் கொள்ளும் கருதுகோள் புள்ளியியல் இலையெனும் கருது

கோளேயே யாகும். இவ் இலேயெனும் கருதுகோள் ஏற்கப்பட்டால் மாதிரி மெண்டீலியன் முழுமைத் தொகுதியைச் சேர்ந்ததெனவும், மறுக்கப்பட்டால் மாதிரி மெண்டீலியன் முழுமைத் தொகுதியைச் சேர்ந்ததெனக் கொள்வதற்குச் சந்தேகம் ஏற்படுகிறது எனவும் கொள்கிறோம். நமது சோதனை அளவை ஒப்பிடப்படும் சிறப்பெல்லேகளுக்குத் தக்கவாறு நூற்றுக்கு ஐந்து, நூற்றுக்கு ஒன்று என்ற முறையில் தவறான முடிவுகளே மேற்கொள்ளும் அபாயம் உள்ளது. நமது கருதுகோள் ஏற்கப்படும்போது, இரு விகிதங்களுக்கான வித்தியாசம் குறிப்பிடத்தக்கபடி அமையாமல் (at 5% not 1%) மாதிரி எடுத்தலினுல் ஏற்பட்ட ஏற்றத்தாழ்வினுல் உண்டானது என்பதாகக் கொள்கிறோம். இதே மறுக்கப்படும்போது, விகிதங்களுக்கான வித்தியாசம் குறிப்பிடும்படி அமைந்து, இம் மாதிரி கொடுக்கப்பட்ட மெண்டீலியன் முழுமைத் தொகுதியைச் சேர்ந்ததெனக் கொள்ள இடமில்லே என்பதாகக் கொள்கிறோம்.

நம்பக எல்லே (Confidence limits):

ஒரு மாதிரியில் உள்ள 1000 பட்டாணிகளேச் சோதனே செய்ததில் 360 பச்சைநிறப் பட்டாணிகள் உள்ளதெனக் கணக்கிடப்பட்டுள்ளது. இம் மாதிரியின் உதவியால் முழுமைத் தொகுதிக்கான பச்சைநிறப் பட்டாணிகளின் நம்பக எல்லேகளேக் கணக்கிடுவோம்.

கொடுக்கப்பட்ட மாதிரியில் பச்சைநிறப் பட்டாணிகள் அமைவதற்கான நிகழ்திறன் 0·36 ஆகவும், மற்ற நிறப் பட்டாணிகள் கிடைப்பதற்கான நிறழ்திறன் 0·64 ஆகவும் உள்ளது. இதனே ஓர் ஈருறுப்புப் பரவலின் மாறியாகக் கொண்டு இதன் திட்டப்பிழை $\sqrt{npq} = \sqrt{1000 \times \cdot 36 \times \cdot 64} = 15 \cdot 18$ எனக் கணக்கிடுகிறோம். 'p'-ன் மதிப்பு எதுவாக அமைந்தாலும்

$$z = \frac{x - np}{\sqrt{npq}}$$

ஓர் இயல்நிலேப் பரவலே ஒத்து அமையும். இயல்நிலேப் பரவலின் தன்மைப்படி $np \pm 1 \cdot 96$ திட்டப்பிழை இடைவெளியில், np-ன் ஒரு மதிப்பு அமைவதற்கான நிகழ்திறன் 0·95 என்றும், $np \pm 2 \cdot 58$ திட்டப்பிழை இடைவெளியில் np அமைவதற்கான நிகழ்திறன் 0·99 என்றும் அறிவோம். நம்மால் கொடுக்கப்படும் எல்லேகளுக்கான நம்பக மதிப்பையே 0·95-ம், 0·99-ம் குறிக்கின்றன. நடைமுறையில் கொடுக்கப்படும் பெரும்பாலான முடிவுகள் 0·95 நம்பக நிகழ்திறனேக் கொண்டதாகும்.

$$|z| = \left|\frac{x-np}{\sqrt{npq}}\right| \leq 1 \cdot 96 = 0 \cdot 95$$

$$|x-np| \leq 1 \cdot 96 \sqrt{npq}$$

$$x - 1 \cdot 96 \sqrt{npq} \leq np \leq x + 1 \cdot 96 \sqrt{npq}$$

கொடுக்கப்பட்ட பச்சைநிறப் பட்டாணிகளின் எண்ணிக்கையிலிருந்து 1·96 திட்டப்பிழை மதிப்பைக் கூட்டியும், குறைத்தும் இம் மாதிரி எடுக்கப்பட்ட முழுமைத் தொகுதிக்கான பச்சை நிறப் பட்டாணிகளின் நம்பக எல்லைகளை அமைக்கிறோம். கொடுக்கப்பட்ட மாதிரியில் உள்ள பச்சைநிறப் பட்டாணிகளின் எண்ணிக்கை 360; அதன் திட்டப்பிழை மதிப்பு 15·18 ஆகும். இம் மாதிரி எடுக்கப்பட்டுள்ள முழுமைத் தொகுதியின் பச்சை நிறப் பட்டாணிகளுக்கான நம்பக எல்லைகள் $360 \pm (1 \cdot 96)(15 \cdot 18)$ அதாவது $360 \pm 29 \cdot 75$ எனக் கொள்கிறோம். இவ் வெல்லைகளில் அமையும் மதிப்புகளுக்கும் நாம் கொடுக்கும் நம்பிக்கையின் நிகழ்திறன் 0·95. நூறு மாதிரிகளில் 95 மாதிரிகளின் மதிப்பு இவ் வெல்லைகளுக்குள் (330·25, 389·75) அமையும் என்பதையே இது குறிக்கும். இவ் வெல்லைகளையே முழுமைத் தொகுதியின் பச்சைநிறப் பட்டாணிகளுக்கான இடைவெளி மதிப்பீடாகக் கொள்கிறோம்.

நிகழ்ச்சிகளின் எண்ணிக்கைக்கு (np) மாறாக நிகழ்ச்சிகள் நிகழும் நிகழ்திறனை (p) அளவையாகக் கொண்டாலும், ஈருறுப்புப் பரவலை உபயோகப்படுத்திச் சிறப்புக்காணும் சோதனைகளை அமைத்துக்கொள்ளலாம். இதன் அளவைகள் p என்ற நிகழ்திறத்தோடும், $\sqrt{pq/n}$ என்ற திட்ட விலக்கத்தோடும் ஓர் இயல் நிலைப் பரவலை ஒத்து அமையும். இதற்கான சோதனை அளவையாக

$$|z| = \left|\frac{p-p_1}{\sqrt{pq/n}}\right|,$$ (p_1 என்பது நிகழ்ச்சிக்கான மாதிரியின் விகித சமம்)

என்ற மாறியின் மதிப்பைக் கணக்கிட்டு 1·96 அல்லது 2·56 உடன் ஒப்பிட்டுக் கருதுகோளை ஏற்றோ மறுத்தோ முடிவு செய்கிறோம்.

காட்டு :

ஒரு வான ஆராய்ச்சி நிலையத்திலிருந்து எடுக்கப்பட்ட குறிப்பில் 900 நாள்களில் 100 நாள்கள் மூடுபனி கவிந்த (foggy) நாள்களாகக் குறிக்கப்பட்டுள்ளன. இவ் வான ஆராய்ச்சி அமைந்த இடத்தின் மூடுபனி கவிந்த நாள்களுக்கான நம்பக எல்லைகளைக் கணக்கிடு.

$$\left.\begin{array}{l}\text{ராண்டமாக எடுக்கப்பட்ட ஒரு நாள் மூடுபனி}\\ \text{கவிந்ததாக இருப்பதற்கான நிகழ்திறன்}\end{array}\right\} = \frac{100}{900} = \frac{1}{9}$$

$$\text{மாதிரியின் பரவலுக்கான திட்டப்பிழை} = \sqrt{pq/n}$$

$$= \sqrt{\frac{1}{9}\cdot\frac{8}{9}\,\frac{1}{900}}$$

$$= 0{\cdot}0105$$

மூடுபனி கவிந்த நாள்களைச் சதவீதமாகக் குறித்தால் 11·11 சத நாள்கள் மூடுபனி கவிந்ததாகவும், அதன் திட்டப் பிழை 1·05 எனவும் அமையும். வான ஆராய்ச்சி நிலையம் அமைந்த இடத்தில் ராண்டமாக எடுக்கப்பட்ட நாள்களுக்கான நம்பக எல்லைகள் $11{\cdot}11 \pm (1{\cdot}96 \times 1{\cdot}05) = 11{\cdot}11 \pm (2{\cdot}06)$ $= (9{\cdot}05 \text{ to } 13{\cdot}17)$ என அமைகின்றன. ராண்டமாக எடுக்கப்பட்ட ஒரு மாதம், வருடத்தில் மூடுபனி கவிந்த நாள்களுக்கான ·95 நிகழ்திற நம்பக எல்லை 9 சதத்திலிருந்து 13 சதம் வரை அமையும் என்பதை இம் மதிப்பீடு காட்டுகிறது.

காட்டு :

மெண்டலியன் கருதுகோட்படி இரு பட்டாணிகளைக் கலக்கச் (cross) செய்வதால் மஞ்சள், பச்சைநிறப் பட்டாணிகளுக்கான விகிதம் 3 : 1 என்ற விகிதத்தில் அமையும் என எதிர்பார்க்கப்படுகிறது. 224 பட்டாணிகள் கொண்ட ஒரு மாதிரியில் 176 மஞ்சள் பட்டாணி இருந்தனவாகக் கணக்கிடப் பட்டுள்ளது. இம் மாதிரி கொடுக்கப்பட்ட மெண்டலியன் கருது கோளுக்கான முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதெனக் கொள்ளமுடியுமா என்பதைச் சோதனை செய்.

மஞ்சள் பட்டாணி கிடைப்பதை 'p' எனக் கொண்டால், 224 பட்டாணிகளில் $224 \times \frac{3}{4} = 168$ பட்டாணிகள் மஞ்சளாக இருக்கவேண்டுமென எதிர்பார்க்கிறோம். 176 பட்டாணிகள் கிடைத்துள்ள நமது மாதிரியில் ஏற்பட்டுள்ள வித்தியாசமானது பட்டாணிகள், மாதிரி எடுத்தலின் ஏற்றத்தாழ்வுகளினால் நிகழ்ந்து மாதிரி அளவை அதிகப்படுத்துவதால் குறையும் எனக் கொள்வதா அல்லது வித்தியாசம் குறிப்பிடும்படி அமைந்து இம் மாதிரி கொடுக்கப்பட்டு மெண்டலியன் முழுமைத் தொகுதியைச் சேர்ந்ததெனக் கொள்ளச் சந்தேகத்திற்கிடமிருக்கிறதா என்பதைச் சோதனை செய்வதே நமது நோக்கமாகும்.

கருதுகோள் : 'மாதிரியில் அமைந்த மஞ்சள்நிறப் பட்டாணிகளுக்கும், மெண்டலியன் முறைப்படி எதிர்பார்க்

கும் மஞ்சள்நிறப் பட்டாணிகளுக்கும் குறிப்பிடத் தக்க வித்தியாசமில்லை.'

செய்முறை :

$$\left.\begin{array}{l}\text{மெண்டலியன் முறைப்படி எதிர்பார்க்கும்}\\ \text{மஞ்சள்நிறப் பட்டாணிகளுக்கான நிகழ்திறன்}\end{array}\right\} = \frac{3}{4} = 0{\cdot}75$$

$$p = 0{\cdot}75$$

$$\left.\begin{array}{l}\text{எடுக்கப்பட்ட மாதிரியில் கிடைத்த மஞ்சள்}\\ \text{நிறப் பட்டாணிகளுக்கான நிகழ்திறன்}\end{array}\right\} = \frac{176}{224} = p_1$$

$$p_1 = 0{\cdot}79$$

$$\text{மாதிரியின் பரவலுக்கான திட்டப்பிழை} = \sqrt{\frac{pq}{n}}$$

$$= \sqrt{\frac{(\cdot 75)\,(\cdot 25)}{224}} = 0{\cdot}029$$

சோதனை அளவை :

$$|z|\left|\frac{p - p_1}{\sqrt{pq/n}}\right|$$

$$= \left|\frac{0{\cdot}75 - 0{\cdot}79}{0{\cdot}029}\right| = 1{\cdot}333$$

$$< 1{\cdot}96$$

சோதனை அளவையின் மதிப்பு 1·96-க்குக் குறைவாக இருப்பதால் முழுமைத் தொகுதிக்கும், மாதிரியில் கிடைத்த மஞ்சள்நிற பட்டாணிகளுக்குமான வித்தியாசம் குறிப்பிடத்தக்கபடி இல்லை என்ற கருதுகோளை ஏற்றுக்கொள்கிறோம். வித்தியாசம் குறிப்பிடத் தக்கபடி அமையாததால், எடுக்கப்பட்ட மாதிரி மஞ்சள், பச்சைநிறப் பட்டாணிகளுக்கான விகிதம் 3 : 1 என்ற மெண்டலியன் முழுமைத் தொகுதியிலிருந்து வந்தது எனக் கொள்கிறோம்.

இரு விகித சமங்களுக்கான வித்தியாசத்தைச் சோதனை செய்தல் :

p_1, p_2 என்ற இரு விகிதசமங்கள் அமைந்த n_1, n_2 உறுப்புகள் கொண்ட இரு மாதிரிகளுக்கான சோதனையைப் பார்ப்போம். இதற்கான சோதனையின் தன்மை, இவ் விரு மாதிரிகளும் ஒரே முழுமைத் தொகுதியைச் சேர்ந்தன அல்லது நாம் குறிப்பிடும் பண்பு ஒரே அளவினதாய் அமைந்த இரு முழுமைத் தொகுதிகளிலிருந்து எடுக்கப்பட்டன என்பதைச் சோதனை செய்வதாக இருக்கும். நமது கருதுகோள் ' இவ் விரு மாதிரிகளிலும் நாம்

குறிப்பிடும் பண்பின் விகிதசமங்களுக்கான வித்தியாசம் குறிப்பிடும்படி இல்லை' என்பதாகும். இக் கருதுகோள் ஏற்கப்படும்போது இரு மாதிரிகளும் ஒரே முழுமைத் தொகுதியைச் சேர்ந்த தென்றோ அல்லது நாம் குறிப்பிடும் பண்பு ஒரே அளவினதாய் உள்ள இரு முழுமைத் தொகுதிகளிலிருந்து எடுக்கப்பட்டது என்றோ கொண்டு, ஏற்பட்ட வித்தியாசம் மாதிரிஎடுத்தலினல் விளைந்த ஏற்றத்தாழ்வினதாகும் என நம்புகிறோம். இதற்கான வித்தியாசம் குறிப்பிடும்படி அமையவில்லை என்பதே நமது முடிவாகும். கருதுகோள் மறுக்கப்படும்போது, வித்தியாசம் குறிப்பிடும்படி அமைந்து இரு மாதிரிகளும் மேற்கண்ட முறையான முழுமைத் தொகுதிகளிலிருந்து எடுத்ததெனக் கொள்வதற்கு இடமளிக்கவில்லை என்ற முடிவைக் கூறுகிறோம்.

நமது கருதுகோள் உண்மையாக இருக்கும்போது, இவ்விரு மாதிரிகளையும் சேர்த்து, முழுமைத் தொகுதியிலுள்ள நமது பண்பிற்கான விகிதத்தை மதிப்பீடு செய்கிறோம்.

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

n_1, n_2 இரு மாதிரிகளின் அளவுகளையும், p_1, p_2 இரு மாதிரிகளிலும் நமது பண்பிற்கான விகிதங்களையும் குறிக்கும்.

இரு மாதிரிகளுக்கான திட்டப்பிழைகள் $\sqrt{\frac{p_1 q_1}{n}}$, $\sqrt{\frac{p_2 q_2}{n_2}}$ என அமைந்து, முழுமைத் தொகுதிக்கான திட்டப்பிழை

$$\sqrt{\frac{pq}{n_1} + \frac{pq}{n_2}} = \sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

என அமைகிறது. இதில் அமையும் p, மேற்சொன்ன முறைப்படி மதிப்பீடு செய்யப்பட்டதாகும்.

சோதனை அளவை :

$$|z| = \left| \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right|$$

ஓர் இயல்நிலைப் பரவலை ஒத்து அமைவதால், இதன் மதிப்பை 1·96, 2·58 உடன் ஒப்பிட்டுக் கருதுகோளை ஏற்றும் மறுத்தும் முடிவு செய்கிறோம்.

காட்டு :

இரு நகரங்களில் முறையே 9743, 39764 பெண்களின் கூந்தல் நிறம் சோதனைக்குள்ளாக்கப்பட்டதில், முதல் நகரத்தில் 4008 பெண்களும், இரண்டாவது நகரத்தில் 17529 பெண்களும் இடைநிலை நிறம்கொண்ட கூந்தலை உடையவர்கள் எனக் கணக்கிடப்பட்டுள்ளது. இவ் விரு மாதிரிகளும், இடைநிலை நிறத்திற்கான விகிதசமம் ஒரே தன்மைத்தான விகிதங்களைப் பெற்ற முழுமைத் தொகுதிகளிலிருந்து எடுக்கப்பட்டன என்பதற்கான சோதனையை அமைப்போம்.

இவ் விரு மாதிரிகளும் ஒரே முழுமைத் தொகுதியைச் சேர்ந்ததாக இருப்பின், இடைநிலை நிறத்திற்கான விகிதசமம் இரு மாதிரிகளிலும் சமமாக இருக்க வேண்டும். $p_1 = p_2$ அல்லது $p_1 - p_2 = 0$ என அமையும். கொடுக்கப்பட்ட மாதிரிகளில் அமைந்துள்ள வித்தியாசம் குறிப்பிடத்தக்கபடி அமைந்துள்ளதா என்பதை அறிவதே நமது சோதனையின் நோக்கமாகும். இதற்காக நாம் அமைக்கும் கருதுகோள் 'இரு மாதிரிகளும் ஒரே முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதென்றோ' அல்லது 'நாம் குறிப்பிடும் பண்பிற்கான விகித சமங்கள் சமமாக உள்ள இரு முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதென்றோ அல்லது இரு மாதிரிகளிலும் கூந்தலின் இடைநிலை நிறத்திற்கான வித்தியாசம் குறிப்பிடும்படி அமையவில்லை' என்றோ இருக்கலாம்.

முதல் மாதிரியில் கூந்தலின் இடைநிலை நிறத்திற்கான விகித சமம் $\Big\} = \dfrac{4008}{9743} = 0{\cdot}411$

$$p_1 = 0{\cdot}411$$

இரண்டாவது மாதிரியில் கூந்தலின் இடைநிலை நிறத்திற்கான விகித சமம் $\Big\} = \dfrac{17529}{39764} = 0{\cdot}441$

$$p_2 = 0{\cdot}441$$

நமது கருதுகோள் உண்மையாக இருக்கும் போது முழுமைத் தொகுதிக்கான கூந்தலின் இடைநிலை நிறத்திற்கான மதிப்பீடு $\Bigg\} = p = \dfrac{n_1p_1 + n_2p_2}{n_1 + n_2}$

$$= \frac{9743 \times {\cdot}411 + 39764 \times {\cdot}441}{9743 + 39744} = 0{\cdot}435$$

மாதிரியின் பரவலுக்கான திட்டப்பிழை $= \sqrt{pq\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$

$$= \sqrt{(0.435)(.565)\left(\frac{1}{9743} + \frac{1}{39765}\right)}$$

$$= 0.056$$

சோதனை அளவை :

$$= |z| = \left| \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right|$$

$$= \left| \frac{.411 - .441}{.056} \right| = \frac{.030}{.056} = 0.535$$

$$= 0.535$$

$$< 1.96$$

சோதனை அளவையின் மதிப்பு 1·96-க்குக் குறைவாக இருப்பதால் நமது கருதுகோள் ஏற்கப்படுகிறது. எனவே, இவ் இரு மாதிரிகளுக்கான கூந்தலின் இடைநிலை நிறத்திற்கான விகிதசமங்களின் வித்தியாசம் குறிப்பிடத்தக்கபடி அமையாமல் இரு மாதிரிகளும் ஒரே முழுமைத் தொகுதியைச் சேர்ந்ததென்ற முடிவை மேற்கொள்கிறோம்.

முழுமைத் தொகுதியின் விகிதசமத்தை மதிப்பிட்டு $\left(p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}\right)$ திட்டப்பிழை $\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ கணக்கிடுவதற்கு மாறாக, p_1, p_2 என்ற மாதிரி அளவுகளைக் கொண்டும், திட்டப்பிழையின் மதிப்பைக் கணக்கிட்டுக் கொள்ளலாம். திட்டப்பிழை அளவு $\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$ எனக் கொள்வது இத்தன்மையைச் சுட்டும். இம் முறையான சோதனை ஒன்றைப் பார்ப்போம்.

காட்டு :

இரு முழுமைத் தொகுதிகளில் சிவப்புநிறப் பூக்கள் கிடைப்பதற்கான விகிதம் 30%, 25% எனக் குறிப்பிடப்பட்டுள்ளது. 1200 மற்றும் 900 விதைகளை எடுத்துச் சோதனை செய்யுமிடத்து சிவப்புநிறப் பூக்கள் கிடைப்பதற்கான விகிதங்களின் வித்தியாசம் குறிப்பிடத்தக்கபடி அமைந்துள்ளதா என்பதைச் சோதனை செய்வோம்.

கருதுகோள் : ‘இரண்டு மாதிரிகளிலும் சிவப்புநிறப் பூக்கள் கிடைப்பதற்கான விகித சமங்களின் வித்தியாசம் குறிப்பிடத்தக்கபடி அமையவில்லை.’

செய்முறை ;

$$\left.\begin{array}{l}\text{முதல் மாதிரியில் சிவப்புநிறப் பூக்கள்}\\ \text{கிடைப்பதற்கான நிகழ்திறன்}\end{array}\right\} = 30\% = \cdot 3 = p_1$$

$$\left.\begin{array}{l}\text{இரண்டாவது மாதிரியில் சிவப்புநிறப்}\\ \text{பூக்கள் கிடைப்பதற்கான நிகழ்திறன்}\end{array}\right\} = \cdot 25\% = \cdot 25 = p_2$$

$$\text{திட்டப்பிழை} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

$$= \sqrt{\frac{(0 \cdot 3) \times (\cdot 7)}{1200} + \frac{(\cdot 25) \times (\cdot 75)}{900}}$$

$$= 0 \cdot 0195$$

சோதனை அளவை:

$$= \left| \frac{p_1 - p_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \right| = \left| \frac{0 \cdot 30 - \cdot 25}{\cdot 0195} \right|$$

$$= 2 \cdot 55$$

சோதனை அளவு 1·96-க்கு அதிகமாகவும், 2·58-க்குக் குறைவாகவும் இருப்பதால், சிறப்புக்காணும் எல்லை 5% ஆக உள்ள போது கருதுகோள் மறுக்கப்பட்டும், சிறப்புக் காணும் எல்லை 1% ஆக உள்ளபோது ஏற்கப்பட்டும் முடிவு செய்கிறோம். சோதனை அளவின் மதிப்பு 2·58-க்கு மிக நெருக்கமாக இருப்பதால், இதற்கான சோதனையை மற்றும் ஒருமுறை செய்வதால் ஓர் உறுதியான முடிவிற்கு வரலாம் என்பதைக் கூறுவது நன்மை பயப்பதாய் உள்ளது.

மாதிரி அளவு (Sample size)

முழுமைத் தொகுதிக்கான விகிதசமத்திற்கும், மாதிரியின் விகிதசமத்திற்கும் உள்ள ஒரு குறிப்பிட்ட அளவு வித்தியாசம் 5% அல்லது 1% சிறப்புக்காணும் எல்லைகளில் குறிப்பிடத்தக்கபடி அமையாதவிடத்து, நாம் எடுக்கக்கூடிய மாதிரியின் அளவு (n) எவ்வளவு இருத்தல் நல்லது என்பதைப் பார்ப்போம்.

முழுமைத் தொகுதிக்கான விகிதசமத்திற்கும், மாதிரியின் விகிதசமத்திற்குமான வித்தியாசம் 0·04 எனக் கொண்டு, இவ் வித்தியாசம் குறிப்பிடத்தக்கபடி அமையாவிடத்து நமது மாதிரியின் அளவு எவ்வளவு இருத்தல் வேண்டும் என்பதைக் கணக்கிடுவோம்.

இவ் வித்தியாசம் குறிப்பிடத்தக்கபடி அமையாவிடத்து நமது சோதனை அளவை 1·96-ம் அதற்குக் குறைவாகவும் அமையும் என்பதை அறிவோம்.

$$|z| = \left|\frac{p-p_1}{\sqrt{pq/n}}\right| < 1{\cdot}96$$

$$p-p_1 = 0{\cdot}04 \text{ (கொடுக்கப்பட்டது)}$$

$$\cdot 04 = 1{\cdot}96\sqrt{pq/n}$$

$$\sqrt{n} = \frac{1{\cdot}96\sqrt{pq}}{\cdot 04}$$

$$\sqrt{n} = 49\sqrt{pq}$$

$$n = 2401\, pq$$

நிகழ்திறன் (p) தெரியாதிருப்பதால் p-ன் மதிப்பை மதிப்பீடு செய்யவேண்டியவர்களாக உள்ளோம். கணக்கியல் முறைப்படி $pq = p(1-p)$ என்பது, $p = \frac{1}{2}$ என அமையுமிடத்துப் பெருமத்தை (maximum) அடைவதைக் காணலாம். p-ன் இம் மதிப்பு உபயோகப்படுத்தும்போது மாதிரியின் அளவிற்கான பெருமம் கிடைக்கப் பெறுவோம்.

$$n = 2401 \times \frac{1}{2} \times \frac{1}{2} = 600$$

நடைமுறையில் மாதிரியின் நிகழ்திற அளவைக்கொண்டு அடியிற்காணும் முறையில் நிகழ்திற அளவு கணக்கிடப்படுகிறது. முழுமைத் தொகுதியிலிருந்து 100 அளவுள்ள மாதிரியை எடுத்து p-ன் மதிப்பை அறிந்து $n = 2401\ pq$ என்ற வாய்பாட்டிலிருந்து கிடைக்கும் மாதிரியின் அளவையை முதல் மதிப்பீடாகக் கொள்கிறோம். இன்னும் சிறப்பான மதிப்பீடு தேவைப்படுமாயின் மதிப்பீடு செய்யப்பட்ட மாதிரி அளவில் பாதியை இரண்டாவது முறையில் எடுத்து, முதலில் எடுத்த அளவுகளையும் சேர்த்துக் கணக்கிடப்படும். p-ன் சிறந்த மதிப்பைக்கொண்டு $n = 2401\ pq$ என்ற வாய்பாட்டிலிருந்து n-ன் மதிப்பை இறுதியாகக் கணக்கிடலாம்.

சிறப்புக்காணும் சோதனைகள் : இயல்நிலைப் பரவல் (Test of significance : Normal distribution)

இதுவரை அமைக்கப்பட்ட சோதனை அளவைகள் யாவும், பண்பின் அடிப்படையில் (countable attribute) பிரிக்கப்பட்டு, சுருறுப்புப் பரவலை ஒத்து அமைவதாகக் கருதப்பட்டு அமைக்கப்பட்டனவாகும். மாதிரியின் அளவுகள் யாவும் அளவுகளின் (measurable) அடிப்படையில் அமைவதாய் இருக்குமிடத்து அமைக்கவேண்டிய சோதனை அளவைகளின் தன்மைபற்றிக்

காண்போம், பண்பின் அடிப்படையில் பிரிக்கப்படும் மாறியின் சோதனை அளவை ஈருறுப்புப் பரவலைத் தழுவியதாகக் கொண்டால், அளவின் அடிப்படையில் பிரிக்கப்படும் மாறியின் சோதனை அளவை இயல்நிலைப் பரவலை ஒத்து அமைவதாகக் கொள்ளப்படும்.

x-என்ற மாறி, μ மற்றும் σ என்ற கூட்டுச் சராசரியையும், திட்டவிலக்கத்தையும் கொண்ட ஓர் இயல்நிலைப் பரவலாக இருப்பின், இம் முழுமைத் தொகுதியிலிருந்து எடுக்கப்படும் மாதிரி, μ என்ற கூட்டுச் சராசரியையும், $\sigma/\sqrt{n}$ என்ற (n = மாதிரியின் அளவு) திட்டப்பிழையையும் கொண்டு ஓர் இயல்நிலைப் பரவலை ஒத்து அமையும்.

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

என்ற மாறி இயல்நிலைப் பரவலை ஒத்து அமைவதால், இதனைச் சோதனை அளவையாகக் கொண்டு, இதன் அளவு 1·96-க்குக் குறைவாக இருப்பின் 5 சதச் சிறப்பெல்லையில் நமது கருதுகோளை ஏற்றும், அதிகமாக இருப்பின் 5 சதச் சிறப்பெல்லையில் நமது கருதுகோளை மறுத்தும் முடிவு செய்கிறோம்.

காட்டு :

900 அளவுகள்கொண்ட ஒரு மாதிரியின் கூட்டுச் சராசரி 3·4 செ.மீ. எனக் கணக்கிடப்பட்டுள்ளது. இம்மாதிரியின் அளவுகள் 3·2 செ.மீ., 2·3 செ.மீ. கூட்டுச் சராசரியையும் திட்ட விலக்கத்தையும் கொண்ட முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதெனக் கொள்ளலாமா என்பதற்கான சோதனையைக் காண்போம்.

நமது கருதுகோள் ‘ முழுமைத் தொகுதியின் கூட்டுச் சராசரிக்கும், மாதிரியின் கூட்டுச் சராசரிக்கும் உள்ள வித்தியாசம் குறிப்பிடும்படி இல்லை ’ என்பதாய் அமைகிறது. ஏற்பட்டுள்ள வித்தியாசம் (3·4 − 3·2 = 2 செ.மீ.) மாதிரிஎடுத்தலினால் ஏற்பட்ட ஏற்ற இறக்கங்களினால் உண்டானதா அல்லது இம் மாதிரியின் அளவுகள் வேறு முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதெனக் கொள்ள இடமுண்டா என்பதே நமது சோதனையாகும்.

மாதிரியின் கூட்டுச் சராசரி = 3·4 செ.மீ.
முழுமைத் தொகுதியின் கூட்டுச் சராசரி = 3·2 செ.மீ.
முழுமைத் தொகுதியின் திட்ட விலக்கம் = 2·3 செ.மீ.

$$\text{திட்டப்பிழை} = \frac{\sigma}{\sqrt{n}} = \frac{2 \cdot 3}{\sqrt{900}} = \frac{2 \cdot 3}{30}$$

$$= 0.077 \text{ செ.மீ.}$$

$$\text{சோதனை அளவை} = Z = \left|\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}\right| = \frac{3 \cdot 4 - 3 \cdot 2}{\cdot 077}$$

$$= \frac{0 \cdot 20}{\cdot 077} = 2 \cdot 6$$

சோதனை அளவையின் மதிப்பு 1·96-க்கும் அதிகமாக இருப்பதால் நமது கருதுகோள் மறுக்கப்பட்டு, முழுமைத் தொகுதியின் கூட்டுச் சராசரிக்கும், மாதிரியின் கூட்டுச் சராசரிக்கும் உள்ள வித்தியாசம் குறிப்பிடும்படி அமைந்து, இம் மாதிரியின் அளவுகள் கொடுக்கப்பட்ட முழுமைத் தொகுதியிலிருந்து ($\mu = 3 \cdot 2, \sigma = 2 \cdot 3$) எடுக்கப்பட்டதெனக் கொள்ள இடமளிக்கவில்லை என்ற முடிவினை மேற்கொள்கிறோம்.

நம்பிக்கை எல்லை, கூட்டுச்சராசரிக்கானது : (Confidence limits, Mean)

n உறுப்புகள்கொண்ட பல மாதிரிகளின் கூட்டுச் சராசரிகள் ஒவ்வொன்றும் முழுமைத் தொகுதியின் கூட்டுச் சராசரிக்கான பலவேறுபட்ட புள்ளி மதிப்புகளைக் கொடுக்கிறதென்றும், இதில் ஒன்றுமட்டும் உண்மையான மதிப்பீடாக இருக்கும் என்றும் குறிப்பிட்டோம். இவ் விடங்களில் மாதிரி ஒன்றிற்கான கூட்டுச் சராசரியிலிருந்து முழுமைத் தொகுதிக்கான கூட்டுச் சராசரியின் நம்பக எல்லைகளைத் தந்து, முழுமைத் தொகுதியின் அளவு இவ் விடைவெளியில் அமைவதற்கான நிகழ்திறத்தையும் கொடுத்து, நமது மதிப்பீட்டிற்கு ஒரு சிறப்பினை ஏற்படுத்துகிறோம்.

$$z = \frac{x - \mu}{\sigma/\sqrt{n}}$$

என்ற அளவை 1·96-க்குக் குறைவாக இருக்கும்போது, இதன் வித்தியாசம் குறிப்பிடத்தக்கதாய் அமையாமல், $\bar{x}$-ன் இருபுறமும் 1·96 திட்டப்பிழை தூரம் 95 சத : அளவுகளைக் கொள்ளும் என்பதை அறிவோம். 0·05 சிறப்பெல்லையாக இருக்கும்போது $-1 \cdot 96\ \sigma/\sqrt{n} < \bar{x} - \mu < 1 \cdot 96\ \sigma/\sqrt{n}$ என்ற எல்லையில் நமது மதிப்பீடு அமைவதற்கான நிகழ்திறன் 0·95

என அமைகிறது. இதனையே $\bar{x} - 1{\cdot}96\ \sigma/\sqrt{n} < \mu < \bar{x} + 1{\cdot}96\ \sigma/\sqrt{n}$ எனவும் குறிக்கலாம். $\bar{x} - 1{\cdot}96\ \sigma/\sqrt{n},\ \bar{x} + 1{\cdot}96\ \sigma/\sqrt{n}$ என்ற இரு எல்லை எண்களே, கூட்டுச் சராசரிக்கான நம்பக வரம்பின் மதிப்புகளாகும்.

காட்டு :

10 வயதுடைய 225 மாணவர்களின் சராசரி எடை 30.கி·கி எனவும் திட்டவிலக்கம் 12 கி.கி. எனவும் உள்ள ஒரு மாதிரியைக் கொண்டு, 10 வயதுடைய மாணவர்களின் முழுமைத் தொகுதிக்கான சராசரி எடையின் நம்பக எல்லைகளை அமைப்போம்.

கூட்டுச் சராசரியின் திட்டப்பிழை $= \sigma/\sqrt{n} = \frac{12}{\sqrt{225}} = \frac{12}{15}$ $0{\cdot}8$ என அமைகிறது. எடுக்கப்பட்ட மாதிரியை ஓர் இயைபிலா மாதிரியாகக் கொண்டால், கூட்டுச் சராசரியின் இருபுறமும் $1{\cdot}96$ திட்டப்பிழை தூரம் 95 சத மாணவர்களின் எடை அமையும் என நம்புகிறோம்.

$$\mu = \bar{x} \pm 1{\cdot}96\ \sigma/\sqrt{n}$$
$$= 30 \pm 1{\cdot}96 \times 0{\cdot}8$$
$$= 30 \pm 1{\cdot}568$$
$$28{\cdot}432 < \mu < 31{\cdot}568$$

10 வயதுடைய மாணவர்களின் முழுமைத் தொகுதிக்கான கூட்டுச் சராசரி எடை $(28{\cdot}432,\ 31{\cdot}560)$ என்ற இடைவெளியில் அமைவதற்கான நிகழ்திறன் $0{\cdot}95$ எனக் குறிப்பிடுகிறோம்.

காட்டு :

20-25 வயதுடையார்தம் சராசரி நாடித்துடிப்பு மணித்துளிக்கு 72ம், திட்டவிலக்கம் மணித்துளிக்கு 9·5 எனவும் கணக்கிடப்பட்டுள்ளது. இம் முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட 55 ஓட்டக்காரர்தம் சராசரி நாடித்துடிப்பு மணித்துளிக்கு 65 என அமைந்தால், ஓட்டக்காரர்தம் நாடித்துடிப்பிற்கும், முழுமைத் தொகுதிக்கானவர்கள் நாடித்துடிப்பிற்கும் உள்ள வித்தியாசம் குறிப்பிடும்படி அமைந்துள்ளதா என்பதைச் சோதனை செய்வோம்.

கருதுகோள்

'20-25 வயதுடையார்தம் நாடித்துடிப்பிற்கும் இதிலிருந்து எடுக்கப்பட்ட 55 ஓட்டக்காரர்கள் நாடித்துடிப்பிற்கும் குறிப்பிடத்தக்க வித்தியாசம் இல்லை'.

முழுமைத் தொகுதிக்கான நாடித்துடிப்பு $= \mu = 72$

ஓட்டர்காரர்கள் நாடித்துடிப்பு $= \bar{x} = 65$

முழுமைத் தொகுதியின் திட்டவிலக்கம் $= \sigma = 9{\cdot}5$

மாதிரியின் பரவலுக்கான திட்டப்பிழை $= \sigma/\sqrt{n} = \dfrac{9{\cdot}5}{\sqrt{55}} = 1{\cdot}28$

$$\text{சோதனை அளவை} = |z| = \left|\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}\right|$$

$$= \frac{65-72}{9{\cdot}5/\sqrt{55}} = \frac{7}{1{\cdot}28} = 5{\cdot}47$$

$$> 2{\cdot}58$$

சோதனை அளவையின் மதிப்பு 2·58-க்கும் அதிகமாக இருப்பதால் நமது கருதுகோள் மறுக்கப்பட்டு, 20-25 வயதுடையார்தம் நாடித் துடிப்பிற்கும் இதிலிருந்து எடுக்கப்பட்ட 55 ஓட்டக்காரர் தம் நாடித்துடிப்பிற்கும் உள்ள வித்தியாசம் குறிப்பிடும்படி அமைந்துள்ளதெனக் கொள்கிறோம்.

இரு கூட்டுச்சராசரிகளின் வித்தியாசத்திற்கான சிறப்புக்காணும் சோதனை (Text of significance for difference between two means).

இரண்டு மாதிரிகளுக்கான கூட்டுச் சராசரியைக் கணக்கிட்டு, இவற்றுக்கான வித்தியாசம் குறிப்பிடத் தக்கபடி அமையாமல், இரண்டும் ஒரே முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதா என்பதைக் காணும் சோதனை அளவையைக் காண்போம்.

n_1, n_2 உறுப்புகள் கொண்ட இரு மாதிரிகள் σ திட்டவிலக்கம் உள்ள முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதாகக் கொள்வோம். n_1, n_2 உறுப்புகள் கொண்ட மாதிரிகள் $\bar{x}_1, \bar{x}_2$ கூட்டுச் சராசரிகளைக் கொடுப்பதாகவும் கொள்வோம். இக் கூட்டுச் சராசரிகளுக்கான வித்தியாசம் குறிப்பிடத்தக்கபடி அமைந்துள்ளதா என்பதே நமது சோதனையாகும்.

$\sigma/\sqrt{n_1}$, $\sigma/\sqrt{n_2}$ என்ற இரு மாதிரிகளின் திட்டப்பிழைகள் இரு மாதிரிகளுக்கான கூட்டுச் சராசரிகளின் வித்தியாசத் திட்டப் பிழையாக $\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}} = \sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ என அமைகிறது. இரு மாதிரிகளும் இயைபிலா மாதிரிகளாக இருக்கும்போது, $(\bar{x}_1 - \bar{x}_2)$ என்ற அளவை பூச்சியக் கூட்டுச் சராசரியையும், $\sigma/\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ என்ற திட்டப்பிழையையும் கொண்டு ஓர் இயல்நிலைப் பரவலை ஒத்து அமையும். இதற்கான சோதனை அளவை $(\bar{x}_1 - \bar{x}_2)/\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$, பூச்சியக் கூட்டுச் சராசரியையும், ஒன்று என்ற திட்ட விலக்கத்தையும் கொண்டு ஓர் இயல்நிலைப் பரவலை ஒத்து அமையும். சோதனை அளவையின் மதிப்பு 1·96-க்குக் குறைவாக இருந்தால் கருதுகோள் ஏற்கப்பட்டும், அதிகமாக இருந்தால் கருதுகோள் மறுக்கப்பட்டும் முடிவுகள் எடுக்கப்படும்.

இரு மாதிரிகளும், திட்டவிலக்கம் தெரிந்த முழுமைத் தொகுதியிலிருந்து எடுக்கப்படாதவிடத்து, மாதிரிகளுக்கான திட்ட விலக்கங்கள் கணக்கிடப்பட்டு, இரண்டு கூட்டுச் சராசரிகளின் வித்தியாசத்தின் திட்டப்பிழை $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ எனக் கொள்ளப்படும்.

காட்டு :

100 பாத்திகளில் கிடைத்த கோதுமையின் சராசரி எடை 200பவு., திட்டவிலக்கம் 10 எனவும், 150 பாத்திகளில் கிடைத்த கோதுமையின் சராசரி எடை 220பவு. எனவும் திட்டவிலக்கம் 12பவு. எனவும் கொடுக்கப்பட்டுள்ளன. முழுமைத் தொகுதியில் கோதுமை விளைவிற்கான திட்டவிலக்கம் 11பவு. எனக் கொண்டு, இரண்டு மாதிரியிலும் கோதுமை விளைவிற்கான வித்தியாசம் குறிப்பிடும்படி உள்ளதா என்பதைச் சோதனை செய்வோம்.

முதல் மாதிரியில் கோதுமை விளைவிற்கான சராசரி மதிப்பு $\} = \bar{x}_1 = 200$

இரண்டாவது மாதிரியில் கோதுமை விளைவிற்கான சராசரி மதிப்பு $\} = \bar{x}_2 = 220$

இவ்விரு மாதிரிகளும் திட்டவிலக்கம் தெரிந்த முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதெனக் கொடுக்கப்பட்டுள்ள தால், இவ் விரு சராசரிகளுக்குமான வித்தியாசத்தின் அளவைக் கான திட்டப்பிழை

$$= \sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= 11\sqrt{\frac{1}{100} + \frac{1}{150}}$$

$$= 1.42$$

$$\text{சோதனை அளவை} = |z| = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{200 - 220}{1{\cdot}42} = 14{\cdot}08$$

$$> 2{\cdot}58$$

சோதனை அளவை 2·58-க்கும் அதிகமாக இருப்பதால், 'இரு மாதிரிகளின் சராசரி விளைவில் குறிப்பிடத்தக்க வித்தியாசம் இல்லை' என்ற நமது கருதுகோள் மறுக்கப்படுகிறது.

காட்டு :

இரு இயைபிலா மாதிரிகளின் அளவுகள் அடியிற்கண்டபடி அமைந்துள்ளது. 6400 நபர்களின் (கோவை மாவட்டம்) சராசரி உயரம் 67·85 அங். திட்டவிலக்கம் 2·5 அங். எனவும் 1600 நபர்களின் (தஞ்சை மாவட்டம்) சராசரி உயரம் 68·55 அங். திட்டவிலக்கம் 2·52 அங். எனவும் கொடுக்கப்பட்டுள்ளது. இரு மாவட்டங்களிலும் உள்ளவர்களின் சராசரி உயரத்திற்கான வித்தியாசம் குறிப்பிடத்தக்கபடி அமைந்துள்ளதா என்பதற் கான சோதனையைக் காண்போம்.

நமது புள்ளியியல் கருதுகோள் 'இரு மாவட்ட மக்களின் சராசரி உயரங்களில் குறிப்பிடத்தக்க வித்தியாசம் இல்லை' என்பதாகும்.

கோவை மாவட்ட நபர்களின் சராசரி உயரம் = 67·85 அங்.

தஞ்சை மாவட்ட நபர்களின் சராசரி உயரம் = 68·55 அங்.

இரு சராசரிகளுக்கான கூட்டுச்சராசரியின் வித்தியாசத்திற்கான அளவுகளின் பரவலுக்கான திட்டப்பிழை $\Big\} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma^2}{n_2}}$

$$= \sqrt{\frac{(2 \cdot 54)^2}{6400} + \frac{(2 \cdot 52)^2}{1600}}$$

$$= 0 \cdot 07 \text{ அங்.}$$

$$\text{சோதனை அளவை} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$|z| = \left|\frac{67 \cdot 85 - 68 \cdot 55}{\cdot 07}\right| = \frac{\cdot 7}{\cdot 07}$$

$$z = 10.$$

$$> 2 \cdot 58$$

இரண்டு மாவட்டங்களுக்கான சராசரி உயரங்களின் வித்தியாசம் குறிப்பிடத்தக்கபடி அமையவில்லை என்ற நமது கருதுகோள் மறுக்கப்பட்டு, இவர்தம் உயரங்களில் வித்தியாசம் இருக்கிறது என்ற முடிவிற்கு வருகிறோம். இச் சோதனையிலிருந்து அதிகக் கூட்டுச் சராசரி உயரம் கொண்ட தஞ்சை மாவட்டத்தினர் கோவை மாவட்டத்தினரைக் காட்டிலும் உயரமானவர்கள் என்று குறிப்பிட இடமுண்டு என்ற முடிவைச் சொல்கிறோம்.

மாதிரி அளவு (Sample size):

ஒரு மாதிரியின் உறுப்புகள் (மாதிரியின் அளவு) எவ்வளவு இருப்பது சிறப்பான மதிப்பீட்டைக் கொடுக்கும் என்பதனைக் காண்போம். முழுமைத் தொகுதியிலுள்ள உறுப்புகள் இயல் நிலைப்பரவலை ஒத்து அமைந்தால், மாதிரியின் 'n' உறுப்புகள் கொண்ட கூட்டுச் சராசரியின் பரவல் $\bar{x}$ என்ற கூட்டுச் சராசரியையும், $\sigma/\sqrt{n}$ என்ற திட்டவிலக்கத்தையும் கொண்டு அமையும். நம்மால் எடுக்கப்படும் முடிவுகளுக்கான சிறப்பெல்லை $\cdot 05$ என உள்ளபோது மாதிரியின் கூட்டுச் சராசரிக்கும் முழுமைத் தொகுதியின் கூட்டுச் சராசரிக்கும் உள்ள வித்தியாசம் இரண்டு திட்டப்பிழை அளவினுக்குக் $[(\bar{x} - \mu) \leq 1 \cdot 96\ \sigma/\sqrt{n}]$ குறைவாக இருப்பது, வித்தியாசம் குறிப்பிடத்தக்கபடி இல்லை என்பதை அறிவோம். இக் கருத்தின் அடிப்படையில்தான் மாதிரியின் உறுப்புகளின் எண்ணிக்கையைக் கணக்கிடுகிறோம்.

காட்டு :

$2 \cdot 2$ திட்டவிலக்கம், N உறுப்புகள் கொண்ட முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட மாதிரியின் கூட்டுச்சராசரிக்கும்

$(\bar{x})$ முழுமைத் தொகுதியின் கூட்டுச் சராசரிக்கும் உள்ள வித்தியாசம் (0·5) குறிப்பிடத்தக்கபடி அமையாவிடத்து, எடுக்கப்பட வேண்டிய 5 சதச் சிறப்பெல்லைக்கான மாதிரியின் அளவைக் கணக்கிடுவோம்.

$(\bar{x}-\mu)=\frac{1}{2}$ என்பது, 5 சதச் சிறப்பெல்லையில் குறிப்பிடத் தக்கபடி அமையாதிருக்கும்போது நமது சோதனை அளவை z 1 96-க்குக் குறைவாக அமையவேண்டும்.

$$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leq 1{\cdot}96$$

$$\bar{x}-\mu = 1{\cdot}96\ \sigma/\sqrt{n}$$

$$\frac{1}{2} = 1\ 96\ \frac{2{\cdot}2}{\sqrt{n}}$$

$$\sqrt{n} = (1\ 96)\ (2{\cdot}2)\ (2) = 8{\cdot}624$$

$$n = (8{\cdot}624)^2$$

$$= 76{\cdot}37$$

$$n = 76$$

மாதிரியின் அளவு $= 76$

காட்டு :

முழுமைத் தொகுதியின் கூட்டுச் சராசரிக்கும் (μ) மாதிரியின் கூட்டுச் சராசரிக்கும் உள்ள வித்தியாசம் திட்ட விலக்கத்தில் 25 சதத்திற்கு அதிகமாக இல்லாமல் இருப்பதற்கான நிகழ்திறன் 0·95 ஆக உள்ள நிலையில், மாதிரிக்கான அளவைக் காண்போம்.

$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leq 1{\cdot}96$ என்பதற்கான நிகழ்திறன் 0·95 என அமையும்.

$$(\bar{x}-\mu) = 1{\cdot}96\ \sigma/\sqrt{n}$$

$$(\bar{x}-\mu) = \sigma/4 \text{ (கொடுக்கப்பட்டது)}$$

$$= 1{\cdot}96\ \times$$

$$\sqrt{n} = 1{\cdot}96\ \sigma \times \frac{4}{\sigma}$$

$$= 1{\cdot}96 \times 4$$

$$n = (1{\cdot}96)^2 \times 16$$

$$= 62$$

மாதிரியின் அளவு $= 62$

பயிற்சி :

1. மஞ்சள், சிவப்புநிறப் பூக்களின் எண்ணிக்கை 3 : 1 என எதிர்பார்க்கப்படும் ஒரு கருதுகோளில் 408, 126 என்றமைந்த மாதிரியின் முடிவுகளை 3 : 1 என்ற விகிதத்திற்கான முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதாகக் கொள்ளலாமா?

2. *A* என்ற நகரத்தில் எடுக்கப்பட்ட 500 பேர் கொண்ட மாதிரியில் 200 பேர் சைவ உணவு உண்பவர்கள் என்றும், *B* என்ற நகரத்தில் எடுக்கப்பட்ட 400 பேர் கொண்ட மாதிரியில் 200 பேர் சைவ உணவு உண்பவர்களாகவும் கணக்கிடப்படுகிறது. இரு நகரங்களிலும் சைவ உணவு உண்பவர்களின் விகிதத்திற்கான வித்தியாசம் குறிப்பிடும்படி அமைந்துள்ளதா என்பதனைச் சோதனை செய்.

3. 400 பழங்கள் கொண்ட ஒரு கூடையில் 40 பழங்கள் பழுக்காமல் இருக்கின்றன எனக் கணக்கிடப்பட்டுள்ளது. இப் பழங்களை ஓர் இயைபிலா மாதிரியாகக் கொண்டு, இப் பழங்கள் எடுக்கப்பட்ட முழுமைத் தொகுதியில் பழுக்காத பழங்களுக்கான நம்பிக்கை எல்லைகளை அமை.

4. ஒரு மாதிரியின் கூட்டுச் சராசரி 100 எனவும் திட்ட விலக்கம் 10 எனவும் கணக்கிடப்பட்டுள்ளது. கொடுக்கப்பட்ட மாதிரியின் கூட்டுச் சராசரிக்கும், எடுக்க உள்ள மாதிரியின் கூட்டுச் சராசரிக்கும் உள்ள வித்தியாசம் 0.01%-க்கு அதிகமாகாமல் இருக்கும்படியான மாதிரியின் அளவைக் கணக்கிடு.

5. ஒரு கூடை முட்டைகளைச் சோதனை செய்ததில் 3427 வெள்ளையாகவும் 522 மாநிறமாகவும் உள்ளன. உயிரியியல் நியதிப்படி 25 சத மாநிற முட்டைகள் எதிர்பார்க்கப்படும் நிலையில் மாதிரியில் உள்ள மாநிற முட்டைகளுக்கான விகித சமத்திற்கும், எதிர்பார்க்கப்படும் விகித சமத்திற்குமான வித்தியாசம் மாதிரி எடுத்தலின் ஏற்ற இறக்கங்களுக்கானதாகக் கொள்ளலாமா என்பதைச் சோதனை செய்.

6. மகப்பேறு மருத்துவமனையில் பிறந்த குழந்தைகளுக்கான (ஓர் ஆண்டில்) பகுப்பில் 3422 ஆண்குழந்தைகள் எனவும்

3121 பெண்குழந்தைகள் எனவும் குறிக்கப்பட்டுள்ளன. பால் விகிதம் 50:50 (sex ratio) என்ற கருதுகோளை இம் முடிவுகளைக் கொண்டு சோதனை செய்.

7. நகராட்சியின் தொடக்கப்பள்ளி ஒன்றில் உள்ள 500 மாணவர்களில் 110 பேர் குறைஇடை (under weight) உள்ளவர்களாகக் கணக்கிடப்படுகிறது. இதனை ஓர் இயைபிலா மாதிரியாகக் கொண்டு, நகராட்சியின் தொடக்கப்பள்ளி மாணவர்களுக்கான குறைஎடை மாணவர்களின் விகிதத்திற்கான நம்பக எல்லைகளை அமை.

8. தமிழ் மொழியில் உள்ள வடமொழிச் சொற்களின் அளவைக் கணக்கிட அடியிற்காணும் சோதனை அமைக்கப்படுகிறது. ஓர் ஆசிரியர் எழுதிய அரசியல் நிர்ணய சபையின் முடிவுகள்பற்றிய குறிப்பில் உள்ள 2025 சொற்களில் 729 வடமொழிச் சொற்களும், அவ் வாசிரியர் எழுதிய அணுச் சக்தி ஆராய்ச்சியின் முடிவுகள்பற்றிய குறிப்பில் உள்ள 1600 சொற்களில் 640 வடமொழிச் சொற்களும் உள்ளன. இவ் வாசிரியர் தம் குறிப்பை ஓர் இயைபிலா மாதிரியாகக் கொண்டு, இவ் வாசிரியர்தம் இரு குறிப்புகளிலும் உள்ள வடமொழிச் சொற்களுக்கான விகித சமங்களின் வித்தியாசம் குறிப்பிடும்படி அமைந்துள்ளதா என்பதற்கான சோதனையைச் செய். தமிழ் மொழியில் வடமொழிச் சொற்கள் அமைவதற்கான நம்பிக்கை எல்லையைக் குறிப்பிடு.

9. 500 அன்னசிப் பழத்தைச் சோதனை செய்ததில் 65 பழங்கள் பழுதானவை எனக் கணக்கிடப்படுகிறது. இம் மாதிரி எடுக்கப்பட்ட முழுமைத் தொகுதியில் பழுக்காத பழத்திற்கான நம்பிக்கை எல்லையைக் குறிப்பிடு. 8·5% முதல் 17·5 % எல்லைக்குள் அமையும் பழுக்காத பழங்கள் கொண்ட கூடைகளின் எண்ணிக்கையைக் கணக்கிடு.

10. ஒரு பள்ளி மாணவர்களின் கூட்டுச் சராசரி எடை 78 பவுண்டுகளாகவும், திட்டவிலக்கம் 5 பவுண்டுகளாகவும் கணக்கிடப்படுகிறது. தற்செயலாக எடுக்கப்பட்ட ஒரு மாணவனின் எடை 69 என அமைகிறது. இவ் எடைக்கும் பள்ளி மாணவர்களின் கூட்டுச் சராசரி எடைக்கும் உள்ள வித்தியாசம் வாய்ப்புக் காரணத்தால் நிகழ்ந்ததெனக் கொள்ளலாமா?

11. ஒரு முழுமைத் தொகுதிக்கான x-மாறியின் கூட்டுச் சராசரி 7·5 எனவும் திட்டவிலக்கம் 20 எனவும், கொடுக்கப்

பட்டுள்ளன. x–ன் 400 அளவுகளைத் தேர்ந்தெடுத்துக் கணக்கிட்டதில் கூட்டுச் சராசரி 7·0 என அமைகிறது. மாதிரியின் கூட்டுச் சராசரிக்கும் முழுமைத் தொகுதியின் கூட்டுச் சராசரிக்கும் உள்ள வேறுபாடு குறிப்பிடும்படி அமைந்துள்ளதா என்பதற்கான சோதனையைச் செய்.

12. 1000, 2000 எண்ணிக்கை உள்ள இரு மாதிரிகளின் கூட்டுச் சராசரி 67·5″, 68·0″ என அமைகிறது. இவ் விரு மாதிரிகளும் திட்டவிலக்கம் 2·5″ உள்ள ஒரு முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதாகக் கொள்ளலாமா ?

13 ஒரு நாட்டில் என்புருக்கி நோயால் பாதிக்கப்பட்டவர்களின் சதவீதம் 0·17 எனக் கொள்ளப்படுகிறது. 500, 1000 மாணவர்கள் உள்ள இரு கல்லூரிகளின் மாணவர்களைச் சோதனை செய்ததில் முறையே 7, 8 மாணவர்கள் நோயால் பாதிக்கப்பட்டிருந்தால் இரு கல்லூரிகளுக்குமான என்புருக்கி நோயினால் பாதிக்கப்பட்டவர்தம் விகிதசமங்களின் வித்தியாசம் குறிப்பிடும்படி உள்ளதா என்பதற்கான சோதனையைச் செய்.

14. ஒரு முழுமைத் தொகுதியின் நாடித்துடிப்பிற்கான கூட்டுச்சராசரி 70 எனவும் திட்டவிலக்கம் 8 எனவும் கணக்கிடப்பட்டுள்ளது. மன நோயால் பாதிக்கப்பட்ட 64 நோயாளிகளின் நாடித்துடிப்பு 75 எனக் கணக்கிடப்பட்டுள்ள மதிப்பு முழுமைத் தொகுதியின் நாடித் துடிப்பிலிருந்து குறிப்பிடும்படியான வித்தியாசம் உள்ளதெனக் கொள்ள முடியுமா?

15. 200 கிராமங்களைக் கொண்ட ஒரு மாதிரியில் கிராமத்திலுள்ள சராசரி நபர்கள் 485 எனவும் திட்டவிலக்கம் 50 எனவும், 200 கிராமங்கள் கொண்ட மற்றொரு மாதிரியில் சராசரி நபர்கள் 510, திட்டவிலக்கம் 40 எனவும் கணக்கிடப்படுகிறது. கிராமங்களின் சராசரி நபர்களுக்கான வித்தியாசம் குறிப்பிடும்படி அமைந்துள்ளதா என்பதைச் சோதனை செய்.

12. t-பரவல்

(t-distribution)

இதுவரை கண்ட சிறப்புக்காணும் சோதனைகள் (Test of significance) யாவும் மாதிரியின் அளவு (sample size) பெரிதாக ($n>30$) இருக்கும்போது சோதனை அளவை (test statistics) ஓர் இயல்நிலைப் பரவலை ஒத்து அமையும் எனக் கொண்டு அமைக்கப்பெற்றவைகளாகும். மாதிரிகளை எடுக்கும் செலவு அதிகமாகுமிடத்தும், மாதிரி உறுப்புகள் நிகழ்வதற்கான நிகழ்திறன் மிக அரிதாக உள்ளபோதும், சோதனைக்குள்ளாக்கும்போது அழியும் தன்மைத்தான மாதிரி உறுப்புகள் உள்ள இடத்தும் மாதிரி அளவு பெரிதாகக் (large samples) கொள்ள இயலாத சில தவிர்க்க முடியாத சூழலில் இருக்கிறோம். மாதிரியின் அளவு குறைந்துள்ளபோது, மாதிரியின் அளவைகளுக்கான பரவலைப் பற்றித் திட்டவட்டமாகத் தெரிந்தால்தான், சிறிய மாதிரிகளுக்கான சோதனையிலிருந்து பொருத்தமான முடிவுகளை மேற்கொள்ள முடியும். சிறிய மாதிரிகள் அமையப்பெற்றவிடத்து அவற்றுக்கான திருத்தமான மாதிரியின் பரவல்களை அமைத்து முடிவினை மேற்கொள்கிறோம். மாதிரியின் அளவு குறைவாக உள்ள இடத்து, இம் மாதிரியின் சோதனை அளவைகள் ஓர் இயல்நிலைப் பரவலை ஒத்து அமையாவிட்டாலும் இவற்றுக்கான தன்மை சில குறிப்பிட்ட திருத்தமான மாதிரியின் பரவல்களாக (Exat sampling distribution) அமைந்துள்ளது குறிப்பிடத்தக்கது. இப்படி அமைந்துள்ள மாதிரியின்பரவல்களில் 't-பரவலும்' ஒன்று. இப் பரவலின் நிகழ்திறத் தன்மையையும், அதன் உதவியால் செய்யக்கூடிய சில முக்கியமான சிறப்புக்காணும் சோதனைகளையும் காண்போம்.

முழுமைத் தொகுதிக்கான கூட்டுச் சராசரி, மாதிரியின் கூட்டுச் சராசரி ஆகியவற்றின் வித்தியாசத்தைக் கூட்டுச் சராசரியின் திட்டப்பிழை அலகினால் அளந்து z என்ற மாறியை

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

அமைத்து, z பூச்சியக் கூட்டுச் சராசரியையும், ஒன்று என்ற திட்டவிலக்கத்தையும் கொண்டு ஓர் இயல்நிலைப் பரவலை ஒத்து அமைந்ததாகக் குறிப்பிட்டோம். 'n' அதிகமாகவும், முழுமைத் தொகுதியின் திட்டவிலக்கம் தெரிந்தும் உள்ள இடத்து இக்கருத்து மிகச் செம்மைப்பட்டுள்ளதை அறிவோம். 'n' குறைவாக உள்ளபோதும், முழுமைத் தொகுதியின் திட்ட விலக்கம் தெரியாத (unknown sigma) இடத்தும், z-என்ற மாறிக்குப் பதிலாக t-என்ற மாறியினை அமைத்து, இயல் நிலைப் பரவலுக்கான நிகழ்திறத்துடன் ஒப்பிடாமல் t-மாறிக்கான பரவலின் நிகழ்திறத்துடன் ஒப்பிட்டுக் கருதுகோளை ஏற்றும் மறுத்தும் முடிவுகளை மேற்கொள்கிறோம்.

$x_1, x_2 \ldots . x_n$ என்ற ஓர் இயைபிலா மாதிரிக்கான உறுப்புகளைக்கொண்டு,

$$\text{கூட்டுச்சராசரி} : \bar{x} = \frac{\Sigma xi}{n}$$

$$\text{திட்டவிலக்கம்} : s = \sqrt{\frac{\Sigma (xi - \bar{x})^2}{n - 1}}$$

ஆகிய அளவைகளைக் கணக்கிட்டு t-என்ற மாறியை அமைக்கிறோம்.

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

μ என்பது முழுமைத் தொகுதியின் கூட்டுச் சராசரியைக் குறிக்கும்.

இம் மாறியின் மாதிரியின் பரவலுக்கான வாய்பாடு முழுமைத் தொகுதியின், திட்டவிலக்கத்தைச் சாராமல் அமைந்திருப்பது t-பரவலின் சிறப்பம்சமாகும். n அதிகமாகவும், முழுமைத் தொகுதியின் திட்டவிலக்கம் தெரிந்த இடத்துமே சிறப்புக்காணும் சோதனைகள் அமைத்த நிலைமாறி—மாதிரியின் உறுப்புகள் குறைவாகவும் முழுமைத் தொகுதியின் திட்ட விலக்கம் தெரியாத இடத்தும் நம்பகமான முடிவுகளை எடுக்க முடியும் என்ற நிலையை உண்டாக்கிய 'ஸ்டூடன்ட்' (student) என்ற புனைபெயரில் எழுதிய W. S. காஸெட் (W. S. Gosset) என்பார்தம் விளக்கமாகிய t-பரவல் புள்ளியியல் துறையில் ஒரு புரட்சியை ஏற்படுத்தியது எனின் மிகையன்று.

இயல்நிலை, ஈருறுப்பு மற்றும் பாய்ஸான் பரவல்களைப் போலவே t-பரவலும் இதற்கெனச் சில தனிப் பண்புகளைப் பெற்று, இப் பரவலைத் தழுவி அல்லது ஒத்து அமையும் மாறிகள் ஒரு குறிப்பிட்ட மதிப்பைப் பெறுவதற்கான நிகழ்திறத்தைக் கொடுப்பதாய் அமைந்துள்ளது. t-பரவலின் அமைப்பு மாதிரி உறுப்புகளின் எண்ணிக்கைக்குத் தக்கவாறு அமைவதால், t-பரவலின் நிகழ்திறன்கள் யாவும் அதன் கட்டின்மை எண்ணிக்கையைக் (degrees of freedom) கொண்டு கணக்கிடப்பட்டு t-பட்டியலாக அமைக்கப்பட்டுள்ளது.

t-பரவலுக்கான நிகழ்திறப் பட்டியல் அடியிற் காணும் முறையில் அமைந்திருக்கும்.

t-பட்டியல்

கட்டின்மை எண்ணிக்கை \ நிகழ்திறன் P	0·9	0·6	0·5	0·4	0·3	0·2	0·1	0·05	0·01
1									
2									
3									
4									
..									
.									
10									
.									
.									
20									
.									
.									
30									

நமது கருதுகோள் உண்மையாக இருக்கும்போது, t-பரவலுக்கான மதிப்புகள் பட்டியலில் கொடுக்கப்பட்டுள்ளன. நமது சோதனை அளவையின் மதிப்பு, அதற்கான கட்டின்மை எண்ணிக்கையையும், நமக்குத் தேவையான சிறப்பெல்லையையும் கொண்டு எடுக்கப்பட்ட t-பட்டியல் மதிப்போடு ஒப்பிடப்படும். பெரும்பாலான இடங்களில் 5 சதம், 1 சதம் சிறப்பெல்லைகளே ஒப்பிடும் தன்மைக்கு எடுத்துக்கொள்ளப்படுவதால், பின்னிணைப்பில் கொடுக்கப்பட்ட t-பட்டியலில் 5%, 1% சிறப்பெல்லைக்கான நிகழ்திற மதிப்புகள் மட்டும் கொடுக்கப்பட்டுள்ளன. கணக்கிடப்பட்ட சோதனை அளவையின் மதிப்பு, பட்டியல் மதிப்பிற்குக் குறைவாக இருப்பின் குறிப்பிட்ட சிறப்பெல்லையில் கருதுகோள் ஏற்கப்பட்டும், அதிகமாக இருப்பின் கருதுகோள் மறுக்கப்பட்டும் முடிவு செய்யப்படும்.

t-பரவலுக்கான கட்டின்மை எண்ணிக்கைபற்றிப் (degrees of freedom) பார்ப்போம். மாதிரியின் கூட்டுச் சராசரி $(\bar{x})$, முழுமைத் தொகுதியின் கூட்டுச் சராசரிக்கு (μ) ஒரு பிறழ்ச்சியற்ற மதிப்பீடாகும் (unbiased estimate) என்பதைக் குறிப்பிட்டோம். முழுமைத் தொகுதியின் திட்டவிலக்கத்திற்கு $(\sigma$ மாதிரியின் திட்டவிலக்கத்தை (s), ஒரு பிறழ்ச்சியற்ற மதிப்பீடாகக் கொள்ளமுடியாதென்பதையும் சுட்டினோம். இதற்கான காரணம், திட்டவிலக்கத்தைக் கணக்கிட உதவும் விலக்கங்கள் யாவும் சார்பற்ற (independent) தன்மையாய் அமையாததே ஆகும். அடுத்த பக்கத்தில் காணும் 10 நபர்களுக்கான எடைகளின் திட்டவிலக்கத்தைக் காண்போம்.

திட்டவிலக்கத்திற்காக அமைக்கப்பட்ட பத்து விலக்கங்களில் முதல் ஒன்பது விலக்கங்களின் கூட்டுத் தொகை $-11 \cdot 1$ ஆகவும், பத்தாவது விலக்கத்திற்கான மதிப்பு $+11 \cdot 1$ ஆகவும் அமைந்து விலக்கங்களின் கூட்டுத் தொகையைப் பூச்சியமாக்குகிறது. கூட்டுச் சராசரியினின்றும் மாறியின் மதிப்புகளுக்காக அமைக்கப்படும் விலக்கங்களின் கூட்டுத் தொகை பூச்சியமாகும் என்ற கூட்டுச் சராசரியின் ஒரு பண்பிற்கேற்ப இது அமைகிறது. ஒன்பது விலக்கங்களின் மதிப்புகள் நிபந்தனை ஏதுமின்றி அமைவதும் பத்தாவது விலக்கத்தின் மதிப்பு மொத்த விலக்கங்களின் கூட்டுத் தொகை பூச்சியமாகும் என்ற நிபந்தனைக்குக் கட்டுப்பட்டு அமைவதையும் உணர்கிறோம். இக் கட்டளை பத்து விலக்கங்களில் எதற்காகவேனும் இருக்கலாம் என்பது உணரத் தக்கது: எந்த விலக்கம் சார்புடையது என்பதைவிட, கொடுக்கப்பட்ட பத்து விலக்கங்களில் ஒன்று தன்னிச்சைப்படி மதிப்

x	$x_i - \bar{x}$
110	−13·9
112	−11·9
115	− 8·9
120	− 3·9
126	+ 2·1
128	+ 4·1
130	+ 6·1
131	+ 7·1
132	+ 8·1
135	+11·1
1239	.

$$\bar{x} = \frac{\Sigma x_i}{n}$$

$$= \frac{1239}{10}$$

$$\bar{x} = 123 \cdot 9$$

பைப் பெறமுடியாது என்பது குறிப்பிடத்தக்கது. σ-ன் மதிப்பாக S ஐக் கொள்வதற்கு இதற்கான திருத்தத்தை ஏற்படுத்துகிறோம்.

n உறுப்புகள்கொண்ட மாதிரியில் $(n-1)$ உறுப்புகளின் விலக்கங்கள் தன்னிச்சைப்படி மதிப்பைப் பெற்றும், ஓர் உறுப்பின் விலக்கம் இவற்றைச் சார்ந்தும் அமைவதால் விலக்கங்களின் கூட்டுத் தொகையை 'n'-க்குப் பதிலாக $(n-1)$ ஆல் வகுத்து S ஐ σ-ன் ஒரு பிறழ்ச்சியற்ற மதிப்பாகக் கொள்கிறோம். t-அளவையில் அமைந்துள்ள S ஐ மதிப்பிடுவதற்கான வரையற்ற விலக்கங்களே t-பரவலின் கட்டின்மை எண்ணிக்கையாகக் (degrees of freedom) கொள்ளப்படுகிறது. கட்டின்மை எண்ணிக்கையின் முழுத் தன்மையையும் உணர உயர்கணித முறைகளை நாட வேண்டி உள்ளதால், இதனை ஒரு சிறு விளக்கமாகக் கொள்ளலாம். கொடுக்கப்பட்ட மாதிரிகளின் அளவிற்குத் தக்கவாறு வரையற்ற விலக்கங்களும் மாறுபடுவதால், கட்டின்மை எண்ணிக்கைக்குத்

தக்கவாறு t-பரவலும் மாறும் தன்மைத்தாகும். இதனால்தான் t-பட்டியலில் கட்டின்மை எண்ணிக்கைக்குத் தக்கவாறு நிகழ் திறன்கள் கொடுக்கப்பட்டுள்ளன. n உறுப்புகள்கொண்ட ஒரு மாதிரிக்கான t-பரவலின் கட்டின்மை எண்ணிக்கை $(n-1)$ ஆகும்.

முழுமைத் தொகுதியின் கூட்டுச் சராசரி தெரிந்தவிடத்து $x_1, x_2 \ldots\ldots x_n$ என்ற அளவுகளைக்கொண்ட மாதிரி கொடுக்கப் பட்ட முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதா என்பதற்கான சோதனையை அமைப்போம்.

மாதிரியின் கூட்டுச் சராசரி $\bar{x}$ எனவும், முழுமைத் தொகு தியின் பிறழ்ச்சியற்ற திட்டவிலக்கம்

$s = \sqrt{\dfrac{\Sigma\ (x_i - \bar{x})^2}{n-1}}$ எனவும் கணக்கிடப்பட்டுள்ளதாகக் கொள்வோம். இவற்றுக்கான t-அளவை அமைத்து, நமது கருதுகோள் உண்மையாக இருக்கும்போது, மாதிரி எடுத்தலின் ஏற்ற இறக்கங்களால் அமையக்கூடிய அதிகப்படியான மதிப்பு களாக அமைந்த t-பட்டியலின் மதிப்போடு ஒப்பிட்டு முடிவு களை மேற்கொள்கிறோம்.

$$|t| = \left|\frac{\bar{x} - \mu}{s/\sqrt{n}}\right| \ (n-1) \text{ கட்டின்மை எண்ணிக்கைக்கானது.}$$

மாதிரி அளவைகளைக்கொண்டு கணக்கிடப்பட்ட சோதனை அளவை, பட்டியலில் $(n-1)$ கட்டின்மை எண்ணிக்கைக்கு ஒரு குறிப்பிட்ட சிறப்பெல்லையில் கொடுக்கப்பட்டுள்ள மதிப்பைவிடக் குறைவாக இருப்பின் நமது கருதுகோள் ஏற்கப்படும். மாதிரியின் கூட்டுச் சராசரிக்கும், முழுமைத் தொகுதியின் கூட்டுச் சராசரிக் கும், உள்ள வேறுபாடு (difference between the population and sample means) குறிப்பிடத்தக்கபடி இல்லை என்பதை ஏற்று, எடுக் கப்பட்ட மாதிரி கொடுக்கப்பட்ட முழுமைத் தொகுதியிலிருந்து வந்ததெனக் கொள்கிறோம். சோதனை அளவையின் மதிப்பு, பட் டியல் மதிப்பைவிட அதிகமாக இருப்பின் மாதிரியின் கூட்டுச் சரா சரிக்கும், முழுமைத் தொகுதியின் கூட்டுச் சராசரிக்கும் உள்ள வித்தியாசம் குறிப்பிடும்படி அமைந்து எடுக்கப்பட்ட மாதிரி கொடுக்கப்பட்ட முழுமைத் தொகுதியிலிருந்து வந்ததெனக் கொள்ள இடமில்லை என்ற முடிவைச் சொல்கிறோம்.

காட்டு :

ஒரு முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட 10 நபர் களின் உயரங்கள் 63, 63, 64, 65, 66, 69, 70, 70 அங். எனக் கணக்-

கிடப்பட்டுள்ளது. இம் மாதிரி 65 அங். உயரம்கொண்ட முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதெனக் கொள்ள முடியுமா என்பதற்கான சோதனையைப் பார்ப்போம்.

மாதிரியின் கூட்டுச் சராசரியையும், பிறழ்ச்சியற்ற திட்ட விலக்க மதிப்பீட்டையும் கணக்கிட்டு t-அளவைக் கணக்கிடுகிறோம். t-பரவலின் கட்டின்மை எண்ணிக்கை ஒன்பதிற்கான 0·05 சிறப்பெல்லையின் பட்டியல் மதிப்போடு ஒப்பிட்டு முடிவை மேற்கொள்கிறோம். கொடுக்கப்பட்ட மாதிரி 65 அங். உயரம் கொண்ட முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதென்ற கருதுகோளை அமைத்துச் சோதனை செய்வோம்.

நபர்களின் எண்ணிக்கை	உயரம் (அங்)	$(x_i - \bar{x})$ d_i	$(x_i - \bar{x})^2$ d_i^2
1	63	—4	16
2	63	—4	16
3	64	—3	9
4	65	—2	4
5	66	—1	1
6	69	+2	4
7	69	+2	4
8	70	+3	9
	70	+3	9
10	71	+4	16
$n = 10$		.	88 $\Sigma(x_i - \bar{x})^2$

$$\text{சராசரி உயரம்} = \frac{670}{10} = 67 \text{ அங்.}$$

$$\text{திட்ட விலக்கம்} = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{88}{9}} = 3{\cdot}13 \text{ அங்.}$$

$$\text{சோதனை அளவு} = t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$= \frac{67 - 65}{3{\cdot}13/\sqrt{10}}$$

$$= \frac{2 \times 3{\cdot}1623}{3{\cdot}13} = 2{\cdot}023$$

t-பட்டியலில் 5 சதச் சிறப்பெல்லையில் கட்டின்மை எண்ணிக்கை ஒன்பதிற்கான மதிப்பு 2·262 என உள்ளது. சோதனை அளவையின் மதிப்புப் பட்டியலில் உள்ள மதிப்பைவிடக் குறைவாக உள்ளதால் நமது கருதுகோள் ஏற்கப்பட்டு, மாதிரியின் சராசரி உயரத்திற்கும் முழுமைத் தொகுதியின் சராசரி உயரத்திற்கும் உள்ள வித்தியாசம் குறிப்பிடும்படி அமையாமல் கொடுக்கப்பட்ட மாதிரியின் அளவுகள் 65 அங். சராசரி உயரங்கொண்ட முழுமைத் தொகுதியைச் சேர்ந்தது என்ற முடிவைச் சொல்கிறோம்.

சிறிய அளவு மாதிரிகளிலும் (small samples) சோதனை அளவையின் உதவியால் முழுமைத் தொகுதியின் அளவைகளுக்கான நம்பக எல்லைகளை (confidence limits) அமைக்க முடியும்.

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

என்பது $(n-1)$ கட்டின்மை எண்ணிக்கைக்கான மதிப்பைக் குறிக்கும். மாதிரியின் கூட்டுச் சராசரிக்கும், முழுமைத் தொகுதியின் கூட்டுச் சராசரிக்கும் உள்ள வித்தியாசம் குறிப்பிடும்படி அமையாத இடத்து, சோதனை அளவையின் மதிப்பு ஒரு குறிப்பிட்ட சிறப்பெல்லையில் பட்டியலில் உள்ள மதிப்பைவிடக் குறைவாக இருக்கும் என்பதை அறிவோம். $t_{.05\,(n-1)}$ என்பது, $(n-1)$ கட்டின்மை எண்ணிக்கைக்கான 5 சதச் சிறப்பெல்லையின் பட்டியல் மதிப்பாகக் கொண்டால்

$$-t_{.05\,(n-1)} < \frac{\bar{x} - \mu}{s/\sqrt{n}} < t_{.05\,(n-1)}$$

$$-t_{.05\,(n-1)} \times \frac{s}{\sqrt{n}} < \bar{x} - \mu < t_{.05\,(n-1)} \times \frac{s}{\sqrt{n}}$$

$$\bar{x} - t_{.05\,(n-1)} \times \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{.05\,(n-1)} \times \frac{s}{\sqrt{n}}$$

நூறு மாதிரிகளில் 95 மாதிரிகளுக்கான மதிப்பு $\left(\bar{x} - t \cdot 05 (n-1) \times \frac{s}{\sqrt{n}}, \bar{x} + t \cdot 05 (n-1) \times \frac{s}{\sqrt{n}}\right)$ என்ற பிரிவுஇடை வெளிக்குள் அமைந்து μ-விற்கான நம்பக இடைவெளியாக அமைகிறது.

காட்டு :

11 எலிகள்கொண்ட மாதிரியின் ரத்தக் குழைமநிலை (viscocity) அளவுகளுக்கான கூட்டுச் சராசரி 3·92 எனவும், திட்டவிலக்கம் 0·61 எனவும் கொடுக்கப்பட்டுள்ளது. மாதிரியின் உதவியால் இவ் வெலிகள் எடுக்கப்பட்ட முழுமைத் தொகுதியில் உள்ள எலிகளின் ரத்தக் குழைம நிலைக்கான கூட்டுச் சராசரியின் நம்பக எல்லைகளை அமைப்போம்.

11 உறுப்புகள்கொண்ட மாதிரியின் 3·92 என்ற கூட்டுச் சராசரியின் அளவிற்கும் முழுமைத் தொகுதிக்கான கூட்டுச் சராசரியின் அளவிற்கும் ஏற்பட்டுள்ள வித்தியாசம் 5 சதச் சிறப்பெல்லையில் குறிப்பிடும்படி அமையாத நிலையில்

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} \leq t \cdot 05 (n-1) \text{ என இருக்கும்.}$$

$$\left[s = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}} \right]$$

$$\bar{x} - t \cdot 05 (n-1) \frac{s}{\sqrt{n}} < \mu < x + t \cdot 05 (n-1) \frac{s}{\sqrt{n}}$$

$$3 \cdot 92 - t \cdot 05 (10) \frac{\cdot 61}{\sqrt{11}} < \mu < x + t \cdot 05 (10) \cdot \frac{\cdot 61}{\sqrt{11}}$$

$$3 \cdot 92 - 2 \cdot 228 \times 0 \cdot 1977 < \mu < x + 2 \cdot 228 \times 0 \cdot 1977$$

[$t \cdot 05 (10) = 2 \cdot 228$ பட்டியல் மதிப்பு]

$$3 \cdot 92 - 0 \cdot 43 \leq \mu \leq 3 \cdot 92 + 0 \cdot 43$$

$$3 \cdot 49 \leq \mu \leq 4 \cdot 35$$

முழுமைத் தொகுதிக்கான ரத்தக் குழைமநிலை கூட்டுச் சராசரியின் நம்பக எல்லைகளாக (3·49, 4·35) அமைகின்றன. ராண்டமாக எடுக்கப்பட்ட ஓர் எலியின் குழைமநிலை அளவு இதில் அமைவதற்கான நிகழ்திறன் 0 95.

n_1 உறுப்புகள்கொண்ட மாதிரியின் கூட்டுச் சராசரி x_1 எனவும் n_2 உறுப்புகள் கொண்ட மாதிரியின் கூட்டுச் சராசரி x_2 எனவும் உள்ள இடத்து, இரண்டு மாதிரிகளுக்குமான கூட்டுச் சராசரிகளின் வித்தியாசம் குறிப்பிடும்படி அமைந்துள்ளதா என்பதற்கான சோதனை.

இதற்கான சோதனை அளவை

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

$(n_1 + n_2 - 2)$ கட்டின்மை எண்ணிக்கையைக்கொண்டு ஒரு t-பரவலைத் தழுவியதாய் அமையும். 's' என்பது இரு மாதிரிகளிலிருந்தும் கணக்கிடப்பட்ட σ-ன் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

$$s^2 = \frac{1}{n_1 + n_2 - 2}\left[\sum_{i:1}^{n_1}(x_{i_1} - \bar{x}_1)^2 + \sum_{i:1}^{n_2}(x_{i_2} - \bar{x}_2)^2\right]$$

இச் சோதனை அளவை t-பரவலை ஒத்து இருப்பதால் நமது கருதுகோள் உண்மையாக இருக்கும்போது, நாம் எதிர்பார்க்கும் $(n_1 + n_2 - 2)$ கட்டின்மை எண்ணிக்கைக்கான ஒரு குறிப்பிட்ட சிறப்பெல்லையின் மதிப்போடு ஒப்பிட்டுக் கருதுகோளை ஏற்றும் மறுத்தும் முடிவு செய்கிறோம்.

காட்டு :

இரண்டு இனங்களைச் சேர்ந்த பன்றிகளில் அதிகப்படியாகக் கிடைத்த எடைகள் கொடுக்கப்பட்டுள்ளன. கொடுக்கப்பட்டுள்ள உணவினுல் ஏற்பட்டுள்ள அதிக எடைகளுக்கான வித்தியாசம் இரண்டு இனங்களிலும் குறிப்பிடும்படி உள்ளதா என்று சோதனையைப் பார்ப்போம்.

பன்றிகளின் வரிசை எண்		1	2	3	4	5	6	7	8
அதிகப்பட்ட எடையின் அளவு	இனம் A	49	53	51	52	47	50	52	53
	இனம் B	52	55	52	53	50	54	54	53

சோதனை அளவையைக் கணக்கிடுவதற்கான கூட்டுச் சராசரி மற்றும் திட்டவிலக்கம் கணக்கிட உதவும் பட்டியலை அமைப்போம்.

x_{1i}	x_{2i}	$x_{1i}-\bar{x}_1$	$(x_{2i}-\bar{x}_2)$	$(x_{1i}-\bar{x}_1)^2$	$(x_{2i}-\bar{x}_2)^2$
49	52	−1·875	−0·875	3·5156	0·7656
53	55	+2·125	+2·125	4·5156	4·5156
51	52	+0·125	−0·875	0·0156	0·7656
52	53	+1·125	+0·125	1·2656	0·0156
47	50	−3·875	−2·875	15·0156	8·7656
50	54	−0·875	+1·125	0·7656	1·2656
52	54	+1·125	+1·125	1·2656	1·2656
53	53	+2·125	+0·125	4·5156	0·0156
407	423			30·8748	17·3748

$$\bar{x}_1 = \frac{407}{8} = 50{\cdot}875$$

$$\bar{x}_2 = \frac{423}{8} = 52{\cdot}875$$

$$s = \sqrt{\frac{1}{n_1+n_2-2}\left[\sum_{i=1}^{n_1}(x_{i1}-\bar{x}_1)^2+\sum_{i=1}^{n_2}(x_{2}-\bar{x}_2)^2\right]}$$

$$= \sqrt{\frac{30{\cdot}8748+17{\cdot}3748}{8+8-2}}$$

$$s = \sqrt{\frac{48{\cdot}2496}{14}} = 1{\cdot}847$$

$$|t| = \left|\frac{\bar{x}_1-\bar{x}_2}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right|$$ (n_1+n_2-2) கட்டின்மை எண்ணிக்கைக்கானது,

$$= \left| \frac{50 \cdot 875 - 52 \cdot 875}{1 \cdot 857 \sqrt{\frac{1}{8} + \frac{1}{8}}} \right| = \frac{2}{1 \cdot 857 \times 0 \cdot 5}$$

$$= 2 \cdot 154$$

5 சதச் சிறப்பெல்லையில், கட்டின்மை எண்ணிக்கை பதினுன்கிற்கான பட்டியல் மதிப்பு 2·145. கணக்கிடப்பட்ட சோதனை அளவு அதிகமாக இருப்பதால் இரண்டு இனப் பன்றிகளின் எடை உயர்விற்கான சராசரி எடையில் உள்ள வித்தியாசம் குறிப்பிடும்படி உள்ளது என்ற முடிவைச் சொல்கிறோம். கொடுக்கப்பட்ட உணவினுல் ஏற்பட்ட எடை உயர்வு ஓர் இனத்தைக் காட்டிலும் மற்றோர் இனத்திற்கு அதிகமாக உள்ளதென்பதைக் காட்டுவதாக முடிவு அமைகிறது.

எடையில் ஏற்பட்ட வித்தியாசம் இரண்டு இனங்களின் தன்மைக்குச் சாதகமானது எனக் கொண்டோம். பன்றிகளின் வயது, உணவு கொடுத்தபோது உள்ள ஆரம்ப எடையின் அளவு முதலியனவும் எடையில் ஏற்பட்ட மாற்றத்தைப் பாதிப்பதாகக் கொள்ளலாம். இச் சூழலில் ஒரே தன்மைத்தான பன்றிகளை (முடிந்தவளவிற்கு) இரண்டிரண்டாக கொண்டு (pair), சோதனை செய்து முடிவிற்கு வருவதே சிறப்பாக உள்ளது. இம் மாதிரி இரண்டிரண்டாக (paired) அமைத்து, இரண்டு மாதிரிகளுக்குமான உறுப்புகள் சமமாக உள்ளபோது சோதனை செய்யும் முறையும் மிகவும் எளிதாக உள்ளது. x_{1i}, x_{2i} என்ற அளவுகளின் வித்தியாசங்களைக் கணக்கிட்டு, இவ் வித்தியாசங்களுக்கான கூட்டுச் சராசரி $(\bar{d})$, திட்டவிலக்கத்தை (s)ஐக் கணக்கிட்டு t-அளவை மதிப்பிடுகிறோம்.

$$t = \frac{\bar{d} - 0}{s/\sqrt{n}}$$

என்ற அளவை ஒரு t-பரவலைச் சார்ந்ததாகக் கொண்டு முடிவு செய்கிறோம்.

காட்டு :

இரு இனங்களைச் சேர்ந்த கோதுமை விளைச்சலுக்கான முடிவுகள் கொடுக்கப்பட்டுள்ளன. இரண்டு இனங்களும் விளைச்சலில் வித்தியாசத்தைக் கணக்கிடுகிறதா என்பதைச் சோதனை செய்வோம்.

சோதனை நிலையம்	1	2	3	4	5	6
முதல் இனம் x_1	24·4	27·9	28·2	19·8	23·1	22·9
இரண்டாவது இனம் x_2	17·5	15·1	21·6	18·2	21·6	13·7
சோதனை நிலையம்	7	8	9	10	11	12
முதல் இனம் x_1	25·6	28·7	26·2	25·7	37·0	31·5
இரண்டாவது இனம் x_2	24·8	27·8	25·2	19·2	34·0	35·2

வ. எ.	X_{1i}	X_{2i}	$X_{1i}-X_{2i}$ d_i	d_i^2
1	24·4	17·5	6·9	47·61
2	27·9	15·1	12·8	163·84
3	28·2	21·6	6·6	43·56
4	19·8	18·2	1·6	2·56
5	23·1	21·6	1·5	2·25
6	22·9	13·7	9 2	84·64
7	25·6	24·8	0·8	0·64
8	28·7	27·8	0·9	0·81
9	26·2	25·2	1·0	1·00
10	25·7	19·2	6·5	42·25
11	37·0	34·0	3·0	9·00
12	31·5	35·2	6·3	39·69
			57·1	437·85

சோதனை அளவைக்கான மதிப்புகள் முன்பக்கத்துக் கீழ்ப் பட்டியலிலிருந்து கணக்கிடப்படுகின்றன.

$$\bar{d} = \frac{\sum d_i}{n} = \frac{57\cdot 1}{12} = 4\ 758$$

$$S_d^2 = \frac{437\cdot 85}{12} - (4\cdot 758)^2$$

$$= 36\cdot 4875 - 22\cdot 638564$$

$$= 13\cdot 8489$$

s-ன் பிறழ்ச்சியற்ற மதிப்பீடு $= \sqrt{\frac{n_1 s^2}{n_1 - 1}} = \sqrt{\frac{13\cdot 8489 \times 12}{11}}$

$$= \sqrt{15\cdot 1078} = 3\ 903$$

$$t = \frac{\bar{d} - 0}{s/\sqrt{n}}$$

$$= \frac{4\cdot 758}{\sqrt{15\cdot 1078}/\sqrt{12}} = \frac{4\cdot 758 \times 3\cdot 4641}{3\cdot 903}$$

$$= 4\cdot 2197 > t\ \cdot 05\ (11) = 2\ 201)$$

கணக்கிடப்பட்ட மதிப்பு எதிர்பார்க்கும் மதிப்பைவிட அதிகமாக இருப்பதால் ($t\ \cdot 05\ (11) = 2\cdot 201$), பன்னிரண்டு சோதனை நிலையங்களில் கிடைத்த இரண்டினங்களுக்கான கோதுமைகளின் சராசரி விளைச்சலில் குறிப்பிடும்படியான வித்தியாசம் உண்டு என்ற முடிவைச் சொல்கிறோம்.

காட்டு :

இரு மாவட்டங்களில் விளைந்த கோதுமையின் புரதச் சத்திற்கான சோதனை முடிவுகள் தரப்பட்டுள்ளன. இரண்டு மாவட்டங்களின் கோதுமைக்கான புரதச் சத்தின் சராசரி மதிப்புகளில் உள்ள வித்தியாசம் குறிப்பிடும்படி அமைந்துள்ளதா என்பதைச் சோதனை செய்வோம்.

மாவட்டம் 1.	12·6	13·4	11·9	12·8	13·0		
மாவட்டம் 2.	13·1	13·4	12·8	13·5	13·3	12·7	12·4

கருதுகோள் : 'இரண்டு மாவட்டங்களில் விளைந்த கோதுமையில் உள்ள புரதச் சத்தின் கூட்டுச் சராசரிகளுக்கான வித்தியாசம் குறிப்பிடும்படி அமையவில்லை.''

செய்முறை :

சோதனை அளவைக்கான மதிப்புகள் அடியிற்காணும் பட்டியலிலிருந்து கணக்கிடப்படுகிறது.

$x_{1}i$	x_{2i}	$x_1 - \bar{x}_2$	$x_{1i} - x_2$	$(x_{1i} - \bar{x}_2)^2$	$(x_{1i} - \bar{x}_2)^2$
12·6	13·1	−0·14	+0·07	0·0196	0.0049
13·4	13·4	+0·66	0·37	0·4356	0·1369
11·9	12·8	−0·84	−0·23	0·7056	0·0529
12·8	13·5	+0·06	0·47	0·0036	0·2209
13·0	13·3	+0·26	0·27	0·0676	0·0729
	12·7		−0·33		0·1089
	12·4		−0·63		0·3969
63·7	91·2			1·2320	0·9943

$$\bar{x}_1 = \frac{63\cdot7}{5} = 12\cdot74$$

$$\bar{x}_2 = \frac{(91\cdot2)}{7} = 13\cdot03$$

$$s^2 = \frac{1\cdot2320 + 0\cdot9943}{5 + 7 - 2} = \frac{2\cdot2263}{10}$$

$$= 0\cdot2226$$

$$s = 0\cdot4719$$

$$|t| = \left| \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right|$$

$$= \left| \frac{12.74 - 13.03}{0.4719 \sqrt{\frac{1}{5} + \frac{1}{7}}} \right|$$

$$t = 1.05$$

5 சதச் சிறப்பெல்லையில், கட்டின்மை எண்ணிக்கை பத்திற்கான பட்டியலின் மதிப்பு

$$t_{.05(10)} = 2.228$$

கணக்கிடப்பட்ட மதிப்பைவிட அதிகமாக இருப்பதால், நமது கருதுகோள் ஏற்கப்படுகிறது. இரண்டு மாவட்டங்களிலும் விளைந்த கோதுமைக்கான புரதச் சத்தில் வித்தியாசம் இல்லை என்ற முடிவை ஏற்றுக்கொள்கிறோம்.

t-பரவலின் உதவியால் உடன்தொடர்பு குழுவிற்கான சிறப்புக் காணும் சோதனை

$(x_1\ y_1)$, $(x_2 y_2)$, ...$(x_n\ y_n)$ என்ற n சோடி உறுப்புகளின் உடன்தொடர்பு கெழுவைக் கணக்கிடுவோம். இவ் வுடன்தொடர்பு கெழு (correlation co-efficient) குறிப்பிடும்படி அமைந்துள்ளதா இல்லையா என்ற சோதனையில் புள்ளியியல் கருதுகோளாக 'இவ் வுறுப்புகள் எடுக்கப்பட்ட முழுமைத் தொகுதியில் இம் மாறிகளுக்கான உடன்தொடர்பு பூச்சியம்' எனக் கொள்வோம். சோதனை அளவையான t-யைக் கணக்கிட்டு, இதன் கட்டின்மை எண்ணிக்கைக்கான நாம் எதிர்பார்க்கும் பட்டியல் மதிப்புடன் ஒப்பிட்டுக் கருதுகோளை ஏற்றும் மறுத்தும் முடிவு செய்கிறோம்.

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

என்ற அளவு $(n - 2)$ கட்டின்மை எண்ணிக்கையுடன் t-பரவலை ஒத்து அமைகிறது. நமது கருதுகோள் உண்மையாக இருக்கும் போது, t-பரவலின் $(n - 2)$ கட்டின்மை எண்ணிக்கைக்கான 0·05 அல்லது 0·01 நிகழ்திற மதிப்போடு சோதனை அளவையை ஒப்பிட்டுக் கருதுகோளை நம்பும் தன்மைக்கான முடிவினைச் செய்கிறோம்.

காட்டு :

18 சோடி உறுப்புகளில் உடன்தொடர்பு கெழு 0·5 எனக் கணக்கிடப்பட்டுள்ளது. இது முழுமைத் தொகுதியில் குறிப்பிடும்படி அமைந்துள்ளதா என்பதைக் காண்போம்.

கருதுகோள் : ' மாதிரியின் உடன்தொடர்பு கெழு (0·5) முழுமைத் தொகுதியில் குறிப்பிடும்படி அமையவில்லை '

முழுமைத் தொகுதியில் இரு மாறிகளும் தொடர்பற்றன என்பதாகக் கொள்கிறோம்.

செய்முறை :

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-\gamma^2}}$$

$$= \frac{0·5 \times \sqrt{18-2}}{\sqrt{1-(0·5)^2}}$$

$$= \frac{0·5 \times 4}{\sqrt{·75}} = \frac{2}{0·865} = 2·31$$

கணக்கிடப்பட்ட அளவு, 16 (18 – 2) கட்டின்மை எண்ணிக்கைக்கான 0·05 நிகழ்திற அளவை $t_{(·05)}\,(16) = 2·12$ விட அதிகமாக இருப்பதால் நமது கருதுகோள் மறுக்கப்படுகிறது. முழுமைத் தொகுதியில் மாறிகள் தொடர்பற்றன என்பது மறுக்கப்பட்டு இரு மாறிகளுக்குமான உடன்தொடர்பு கெழு குறிப்பிடும்படி அமைந்துள்ளது எனக் கொள்கிறோம்.

உடன்தொடர்பு கெழுவிற்கான பரவல் இயல்நிலைப் பரவலை ஒத்து அமையாமல் சமச்சீரற்ற தன்மையாய் அமைந்துள்ளதால் அதற்கான திட்டப்பிழையை உபயோகப்படுத்திச் செய்யும் சிறப்புக்காணும் சோதனைகள் முற்றிலும் நம்பும்படியான முடிவுகளை மேற்கொள்ள ஏதுவாய் அமையாமல் உள்ளன. 'ρ' என்பது முழுமைத் தொகுதிக்கான உடன்தொடர்பு கெழுவாகவும் 'γ' என்பது மாதிரியின் உடன்தொடர்பு கெழுவாகவும் உள்ளவிடத்து, ஃபிஷர் (Fisher), ' z' என்ற மாறியை அமைத்து, இவற்றுக்கான வித்தியாசம் பூச்சியச் சராசரியையும் $\frac{1}{\sqrt{n-3}}$ என்ற திட்டவிலக்கத்தையும் கொண்டு ஓர் இயல்நிலைப் பரவலை ஒத்து அமைகிறது எனக் குறிப்பிட்டுள்ளார். இதன் உதவியால் முழுமைத் தொகுதிக்கும், மாதிரிக்கும் அல்லது இரு மாதிரிகளுக்குமான உடன்தொடர்பு கெழுக்களுக்கான சிறப்புக்காணும் சோதனைகளை அமைக்கலாம்.

காட்டு :

28 சோடி உறுப்புகளுக்கான உடன்தொடர்பு கெழு 0·68 என உள்ளவிடத்து, இம் மாதிரியின் உறுப்புகள் முழுமைத் தொகுதியில் சார்பற்ற தன்மைத்து என்பதற்கான சோதனையைச் செய்வோம்.

முழுமைத் தொகுதியில் இம் மாறிகள் தொடர்பற்றன என்பதால் $\rho = 0$ என அமைகிறது. $\gamma = 0\cdot 68$ என்பதை ஃபிஷரின் z-மாறியாக மாற்றியமைக்க வேண்டும்.

$$z = \frac{1}{2} \log_e \frac{1+\gamma}{1-\gamma}$$

$$= 1\cdot 1513 \log_{10} \frac{1+\gamma}{1-\gamma}$$

$$= 1\cdot 1513 \times \log \frac{1\cdot 68}{0\cdot 32}$$

$$= 0\cdot 829$$

$$\rho = 0$$

$$\text{சோதனை அளவு} = \frac{z_1 - 0}{1/\sqrt{n-3}}$$

$$= \frac{0\cdot 829 - 0}{1/\sqrt{25}} = 0\cdot 829 \times 5$$

$$= 4\cdot 145$$

$$> 1\cdot 96$$

முழுமைத் தொகுதியில் இம் மாதிரிக்கான தொடர்பற்ற தன்மை மறுக்கப்பட்டு, மாதிரியின் உடன்தொடர்பு கெழு குறிப்பிடும்படி அமைந்துள்ளது என்ற முடிவைச் சொல்கிறோம்.

n_1, n_2 அளவுகள் கொண்ட இரு மாதிரிகளின் உடன்தொடர்பு கெழுக்களுக்கான (γ_1, γ_2) வித்தியாசம் குறிப்பிடும்படி உள்ளதா இல்லையா என்பதற்கான சோதனை

$$z_1 = \frac{1}{2} \log_e \frac{1+\gamma_1}{1-\gamma_1}$$

$$z_2 = \frac{1}{2} \log_e \frac{1+\gamma_2}{1-\gamma_2}$$

z_1-மாறியின் திட்டப்பிழை $\frac{1}{\sqrt{n_1 - 3}}$ எனவும், z_2 மாறியின் திட்டப்

பிழை $\frac{1}{\sqrt{n_1-3}}$ எனவும் அமைவதால் (z_1-z_2) மாறிக்கான திட்டப் பிழை $\sqrt{\frac{1}{n_2-3}+\frac{1}{n_2-3}}$ என அமைகிறது.

சோதனை அளவை ;

$$\frac{z_1 - z_2}{\sqrt{\frac{1}{n_1-3}+\frac{1}{n_2-3}}}$$

எனக்கொண்டு, இதன் மதிப்பு 1·96-க்கு அதிகமாயின் கருது கோள் மறுக்கப்பட்டும், 1·96-க்குக் குறைவாயின் கருதுகோள் ஏற்கப்பட்டும் முடிவுகள் செய்கிறோம்.

காட்டு :

420 மாணவர்களின் எடை, உயரத்திற்கான உடன்தொடர்பு கெழு 0·72 ஆகவும், 225 மாணவர்களின் எடை உயரத்திற்கான உடன்தொடர்பு கெழு 0·65 எனவும் உள்ள இரு மாதிரிகளின் உடன்தொடர்பு கெழுக்களுக்கான வித்தியாசம் குறிப்பிடும்படி அமைந்துள்ளதா என்பதைச் சோதனை செய்வோம்.

கொடுக்கப்பட்ட அளவுகள் :

$$n_1 = 420 \qquad r_1 = 0{\cdot}72$$
$$n_2 = 225 \qquad r_2 = 0{\cdot}65$$

கருதுகோள் : ‘இரு மாதிரிகளுக்குமான உடன்தொடர்பு கெழுக்களின் வித்தியாசம் குறிப்பிடும்படி இல்லை.’

செய்முறை :

$$z_1 = \frac{1}{2}\log_e \frac{1+\gamma_1}{1-\gamma_1}$$
$$= \frac{1}{2}\log_{10} \frac{1+\gamma_1}{1-\gamma_1} \times 1{\cdot}1513$$
$$= \frac{1}{2}\log_{10} \frac{1{\cdot}72}{1-72} \times 1{\cdot}1513$$
$$= 0{\cdot}89$$
$$z_2 = \frac{1}{2}\log_e \frac{1+\gamma_2}{1-\gamma_2}$$

$$= \frac{1}{2} \log_{10} \frac{1+\gamma_2}{1-\gamma_2} \times 1 \cdot 1513$$

$$= 1 \cdot 1513 \times \frac{1}{2} \log \frac{1 \cdot 65}{1 - \cdot 65} = 0 \cdot 77$$

திட்டப்பிழை $\sigma z_1 - z_2 = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}$

$$= \sqrt{\frac{1}{417} + \frac{1}{222}} = 0 \cdot 08308$$

சோதனை அளவை $= \dfrac{z_1 - z_2}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}} = \dfrac{0 \cdot 89 - 0 \cdot 77}{0 \cdot 08308}$

$$= \frac{0 \cdot 12}{0 \cdot 08308} = 1 \cdot 4443$$

$$< 1 \cdot 96$$

கணக்கிடப்பட்ட மதிப்பு 1·96 ஐ விடக் குறைவாக இருப்பதால், இரண்டு மாதிரிகளின் உடன்தொடர்பு கெழுக்களுக்கான வித்தியாசம் குறிப்பிடும்படி இல்லை எனக் கொண்டு இரண்டு மாதிரிகளும் ஒரே முழுமைத் தொகுதியைச் சேர்ந்ததெனச் சொல்கிறோம்.

மாறிகளின் தொடர்பு கெழுவிற்கான (Regression co-efficient) சிறப்புக்காணும் சோதனை

$(x_1\ y_1,), (x_2\ y_2) \ldots\ldots\ldots (x_n\ y_n)$ என்ற n உறுப்புகள் கொண்ட மாதிரி கொடுக்கப்பட்டுள்ளதாகக் கொள்வோம். மாறிகளின் தொடர்பு சமன்பாட்டிற்கான தன்மை (x-தனித்த மாறியாக உள்ளபோது)

$$(y_i - \bar{y}) = b_1 (x_i - \bar{x}) \text{ என அமையும்.}$$

$$b_1 = \frac{\mu_{11}}{\sigma x^2}$$

x தனித்த மாறியாக இருக்கும்போது y-ன் சிறந்த மதிப்பீட்டை

$$y_i = \bar{y} + b_1 (x_i - \bar{x})$$

என்ற சமன்பாட்டிலிருந்து பெறலாம். முழுமைத் தொகுதியில் x–தனித்த மாறியாக உள்ளபோது, B_1 (முழுமைத் தொகுதியின் y-மாறியை x-மாறியில் குறிப்பிடும் மாறியின் தொடர்பு கெழு : (Regression co-efficient of y or x) என்ற தொடர்பு கெழு

இருப்பதாகக் கொள்ளின், மாதிரியின் மாறியின் தொடர்பு கெழுவிற்கும் (b_1), முழுமைத் தொகுதியின் மாறியின் தொடர்பு கெழுவிற்கும் (B_1) குறிப்பிடும்படியான வித்தியாசம் இருக்கிறதா என்பதே நமது சோதனையாகும்

நமது கருதுகோள் உண்மையாக இருக்கும்போது

$$t = (b_1 - B_1) \sqrt{\frac{(n-2)\,\Sigma\,(x_i - \bar{x})^2}{\Sigma\,(y_i - \bar{y})^2}}$$

என்ற மாறி $(n - 2)$ கட்டின்மை எண்ணிக்கையுடன் t-பரவலை ஒத்து அமைகிறது எனக் கொண்டு, குறிப்பிட்ட சிறப்பெல்லையில் எதிர்பார்க்கப்படும் மதிப்புடன் கணக்கிடப்பட்ட மதிப்பு ஒப்பிடப் பட்டுக் கருதுகோளை மறுத்தும், ஏற்றும் முடிவு செய்யப்படும்.

காட்டு :

$y = 15{\cdot}13 - 0{\cdot}1608\ x$ என்ற சார்பு சமன்பாட்டின் சரிவு (slope) பூச்சியத்திற்கு மாறுபட்டுள்ளதா என்பதற்கான சோதனையைச் செய்வோம். $n = 40$, $\Sigma\,(y_i - \bar{y})^2 = 2260$, $\Sigma\,(x_i - \bar{x}) = 2600$ என்பன கொடுக்கப்பட்ட மதிப்புகளாகும்.

$b = 0$ என்ற சோதனை $\frac{\mu_{11}}{\sigma x^2} = 0$ என்பதற்குச் சமமாதலின் கருதுகோள் ஏற்கப்படும்பொழுது, முழுமைத் தொகுதியில் இரு மாறிகளுக்கும் உடன்தொடர்பு இல்லை என்ற முடிவைக் காட்டும்.

$$t = 0{\cdot}1608\sqrt{\frac{38\ (2600)}{2260}}$$

$$= 0{\cdot}1608 \times 6{\cdot}537$$

$$= 1{\cdot}051$$

$$< t_{(\cdot 05)\,38)} = 1{\cdot}96$$

கணக்கிடப்பட்ட மதிப்பு எதிர்பார்க்கும் மதிப்பைவிடக் குறை வாக இருப்பதால், $b_1 = 0$ எனும் நமது கருதுகோள் ஏற்கப்பட்டு முழுமைத் தொகுதியில் இரு மாறிகளும் தொடர்பற்றுள்ளன என்ற முடிவைச் சொல்கிறோம்.

காட்டு :

$y = 1{\cdot}12x - 1{\cdot}79$ என்ற தொடர்பு சமன்பாட்டிற்கான மாறிகளின் தொடர்பு கெழு, முழுமைத் தொகுதியின் மாறிகளின்

தொடர்பு கெழு ஒன்று எனக் கொள்ள முடியுமா என்பதைச் சோதனை செய்வோம். $n = 11$, $\Sigma(x_i - \bar{x})^2 = 301826$, $\Sigma(y_i - \bar{y})^2 = 2766$ எனவும் கொடுக்கப்பட்டுள்ளன.

'$y = 1 \cdot 12x - 1 \cdot 79$ என்ற சமன்பாட்டிற்கான தொடர்பு கெழு ($1 \cdot 12$), முழுமைத் தொகுதியின் தொடர்பு கெழுவான ஒன்றிலிருந்து குறிப்பிடும்படியான வித்தியாசத்தைப் பெற்றுள்ளது' என்பது நமது கருதுகோளாகும்.

$$n = 11, \; b_1 = 1 \cdot 12, \; B_1 = 1 \cdot 0$$

$$\Sigma(x_i - \bar{x})^2 = 301826 \quad \Sigma(y_i - \bar{y}) = 2766$$

$$t = (1 \cdot 12 - 1 \cdot 0 \sqrt{\frac{9\ (301826)}{2766}} = 3 \cdot 76$$

கட்டின்மை எண்ணிக்கை ஒன்பதிற்கான, ·05 சிறப்பெல்லையில் நாம் எதிர்பார்க்கும் t–பரவலின் மதிப்பான $2 \cdot 26$ ($t_{\cdot 05}(9) = 2 \cdot 26$) ஐ விடக் கணக்கிடப்பட்ட மதிப்பு அதிகமாக இருப்பதால் நமது கருதுகோள் மறுக்கப்பட்டு, இரண்டிற்குமான வித்தியாசம் குறிப்பிடும்படி அமைந்துள்ளது என்ற முடிவைச் சொல்கிறோம்.

மாறிகளின் தொடர்பு கெழுவிற்கான நம்பக எல்லைகள்

மாறிகளின் கொடுக்கப்பட்ட சமன்பாட்டிற்கான தொடர்பு கெழு (b), முழுமைத் தொகுதிக்கான தொடர்பு கெழு (B)ஐ விடக் குறிப்பிடும்படியான வித்தியாசத்தைப் பெறாமல் இருக்கும்போது

$$|t| = \frac{b - B}{SE_b} = (b - B)\sqrt{\frac{(n-2)\,\Sigma(x_i - \bar{x}^2)}{(y_i - \bar{y})^2}}$$

[SE_b = standard error of b]

$$\leq t_{\cdot 05}\,(n-2)$$

$$B = b \pm t_{\cdot 05}(n-2)\sqrt{\frac{(n-2)\,\Sigma(x_i - \bar{x})^2}{\Sigma(y_i - \bar{y})^2}}$$

$$b - t_{\cdot 05}\,(n-2)\sqrt{\frac{(n-2)\,\Sigma(x_i - \bar{x})^2}{\Sigma(y_i - \bar{y})^2}} \leq B$$

$$\leq b + t_{\cdot 05}(n-2)\sqrt{\frac{(n-2)\,\Sigma(x_i - \bar{x})^2}{\Sigma(y_i - \bar{y})^2}}$$

காட்டு:

$y = 12{\cdot}13 - 0{\cdot}1608\ x$ என்ற தொடர்பு சமன்பாட்டிற்கான மாறிகளின் தொடர்பு கெழுவின் நம்பக எல்லைகளைக் காண்போம். $n = 24$, $\Sigma(x_i - \bar{x})^2 = 1300$, $\Sigma(y_i - \bar{y})^2 = 1160$ எனவும் கொடுக்கப்பட்டுள்ளது.

$$B = b \pm t_{\cdot 05\,(n-2)} \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{(n-2)\,\Sigma(x_i - \bar{x})^2}}$$

$$B = -\,0.16 \pm 2{\cdot}074\sqrt{\frac{1160}{22 \times 1130}}$$

$$= -\,0{\cdot}16 \pm 2{\cdot}074 \times {\cdot}202$$

$$= -\,0{\cdot}16 \pm 0{\cdot}42$$

95 சத நம்பக எல்லைகள் $-\,0{\cdot}58 < B < +\,0{\cdot}26$

100 மாதிரிகளில், 95 மாதிரிகளில் மாறிகளின் தொடர்பு கெழுக்களின் மதிப்புகள் $(-\,0{\cdot}58, +\,0{\cdot}26)$ இடைவெளிக்குள் அமைவதற்கான நிகழ்திறன் 0·95 எனவும் குறிப்பிடலாம்.

பயிற்சி :

1. மாதிரியின் அளவு 30 உள்ள ஒரு மாதிரியின் உடன் தொடர்பு கெழு 0·75 எனக் கொடுக்கப்பட்டுள்ளது. உடன் தொடர்பு கெழு 0·90 என உள்ள முழுமைத் தொகுதியிலிருந்து இம் மாதிரி எடுக்கப்பட்டுள்ளதா என்பதற்கான சோதனையைச் செய்.

2. 23 (n_1) உறுப்புகள் கொண்ட மாதிரியின் உடன்தொடர்பு கெழு 0·5 எனவும் 28 (n_2) உறுப்புகள் கொண்ட மாதிரியின் உடன்தொடர்பு கெழு 0·8 எனவும் கொடுக்கப்பட்டுள்ளது. இவ் விரண்டு உடன்தொடர்பு கெழுக்களுக்கான வித்தியாசம் குறிப்பிடும்படியாக அமைந்துள்ளதா என்பதற்கான சோதனையை அமை.

3. 28 சோடி அளவுகளுக்கான உடன்தொடர்பு கெழு 0·7 எனக் கொடுக்கப்பட்டுள்ளது. முழுமைத் தொகுதியின் உடன் தொடர்பு கெழு 0·57 என உள்ள முழுமைத் தொகுதியிலிருந்து 28 சோடி உறுப்புகளும் எடுக்கப்பட்டதெனக் கொள்ளலாமா? உடன்தொடர்பு கெழுவிற்கான 95% நம்பக எல்லைகளை அமை

4. 10, 12 சோடி உறுப்புகள்கொண்ட இரு மாதிரிகளுக்கான உடன்தொடர்பு கெழுக்கள் 0·8, 0·6 எனக் கொடுக்கப்பட்டுள்ளன. இவ் விரண்டு மாதிரிகளும் ஒரே முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டவை என்பதற்கான சோதனையைக் காண்.

5. 63, 80 சோடி மாணவர்கள் கொண்ட இரு மாதிரிகளின் அறிவுக் குறியீட்டெண் (I. Q.), கணிப்புத்திறம் (Arithmetic) ஆகியவற்றின் உடன்தொடர்பு கெழுக்கள் முறையே 0·63, 0·55 எனக் கொடுக்கப்பட்டுள்ளன. இரண்டு தொகுதிகளுக்குமான அறிவுக் குறியீட்டெண், கணிப்புத்திறம் ஆகியவற்றின் உடன் தொடர்பு கெழுக்களுக்கான வித்தியாசம் குறிப்பிடும்படி உள்ளதா என்பதற்கான சோதனையைச் செய்.

6. மாதிரியின் அளவு 52 என உள்ள மாதிரியின் உடன் தொடர்பு கெழு 0·65 எனக் கணக்கிடப்படுகிறது. இம் மாதிரி எடுக்கப்பட்ட முழுமைத் தொகுதிக்கான உடன்தொடர்பு கெழு 0·5 என இருக்க இடமுண்டா என்பதற்கான சோதனையைச் செய். முழுமைத் தொகுதிக்கான உடன்தொடர்பு கெழுவின் 99 சத நம்பக எல்லைகளை அமை.

7. 12 நோயாளிகளுக்கு ஒருவகையான கிளர்ச்சியூட்டும் மருந்து (stimulus) கொடுத்து, அவர்களின் ரத்த அழுத்தத்தில் ஏற்பட்ட மாறுதல்கள் குறிக்கப்பட்டுள்ளன. 5, 2, 8, −1, 3, 0, 6, −2, 1, 5, 0, 4. ‘இக் கிளர்ச்சியூட்டி கொடுக்கும்போதெல்லாம், ரத்த அழுத்தமும் குறிப்பிடும்படி அதிகமாகும்’ என்ற கருதுகோளின் முடிவிற்கான சோதனையைக் கொடுக்கப்பட்ட ரத்த அழுத்த உயர் அளவுகளைக் கொண்டு அமை.

8. கோதுமைச் செடிகள் முளைக்கும் தன்மையைக் காண *A*, *B* என்ற சோதனை முறைகள் கையாளப்பட்டன. *A* என்ற சோதனைக்குள்ளாக்கப்பட்ட 9 பாத்திகளில் கிடைத்த முளைத் தோன்றல்களின் எண்ணிக்கைகளும், *B* என்ற சோதனைக்குள்ளாக்கப்பட்ட 8 பாத்திகளில் கிடைத்த முளைத் தோன்றல்களின் எண்ணிக்கைகளும் கொடுக்கப்பட்டுள்ளன.

A	17	26	18	25	27	28	26	23	17
B	16	16	22	21	18	15	16	20	—

A என்ற சோதனையில் கிடைத்த முளைகள் குறிப்பிடும்படி அதிகமாக உள்ளதா என்பதற்கான சோதனையைச் செய்.

9. எலும்பு வளர்ச்சிக்காக எட்டுச் சோடி பன்றிகட்கும் சுண்ணும்புச்சத்து, எலும்புச்சத்துக் கலந்த இருவகை உணவுகள் கொடுக்கப்பட்டுள்ளன. உணவுகளின் தரத்தைக் சோதிக்கச் செய்யப்பட்ட இச் சோதனையில் அடியிற்காணும் முடிவுகள் பெறப் பட்டன. இருவகை உணவுகளின் கூட்டுச் சராசரிகளுக்கான வேற்றுமை குறிப்பிடும்படி உள்ளதா என்பதைச் சோதனை செய்.

சோடி எண்	சுண்ணும்புச் சத்து	எலும்புச் சத்து
1	49·2	51·5
2	53·3	54·9
3	50 6	52·2
4	52·0	53·3
5	46·8	51·6
6	50·8	54·1
7	52·1	54·2
8	53·0	53·3

10, ஒரு குறிப்பிட்ட முழுமைத் தொகுதியிலிருந்து 10 நபர்கள் கொண்ட மாதிரியை எடுத்ததில், அவர்களுக்கான உயரம் 61, 65, 66, 67, 68, 70, 70, 70, 71, 72 என அமைகிறது. இவர்களுக்கான சராசரி உயரம் 66 அங். எனக் கொள்ளமுடியுமா என்பதற்கான சோதனையைச் செய்.

11, ஒரு மருத்துவமனையில் 10 நோயாளிகளுக்கு இரு வகைத் தூக்க மருந்துகளைக் கொடுத்து, அவர்கள் பொதுவாகத் தூங்கும் நேரங்களில் ஏற்பட்ட மாற்றங்கள் பின்வருமாறு:

நோயாளி		1	2	3	4	5	6	7	8	9	10
தூங்கும் நேரத்தில் ஏற்பட்ட மாற்றம்	ம₁	+·7	−1·6	−·2	−1·2	−·1	3·4	3·7	−·8	0	2·0
	ம₂	+1·9	0·8	1 1	+·1	−·1	4·4	5 5	1 6	4·6	3·4

12. A என்ற உணவு 12 பிராணிகளுக்கும், B என்ற உணவு 15 பிராணிகளுக்கும் கொடுத்த சோதனையில் கிடைத்த எடை உயர்விற்கான முடிவுகள் கொடுக்கப்பட்டுள்ளன.

A	25	32	30	34	24	25	14	32	24	30
B	44	34	22	10	47	31	40	30	32	35
A	31	35								
B	18	21	35	29	22					

எடை உயர்வில் இரண்டுவகை உணவுகளுக்குமான சராசரி அளவில் குறிப்பிடும்படியான வித்தியாசம் உள்ளதா என்பதை சோதனை செய்.

13. $y = 5 - 0{\cdot}2x$ என்ற தொடர்பு சமன்பாட்டின் மாறிகளின் தொடர்பு கெழு, '0' என்ற 'மாறிகளின் தொடர்பு கெழு' கொண்ட முழுமைத் தொகுதியிலிருந்து அமைந்ததாகக் கொள்ளலாமா ?

14. x எனும் மாறி 1, 2, 3, ... 30 எனும் மதிப்புகளைப் பெறுமிடத்து $\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 599{\cdot}62$, $\Sigma(y_i - \bar{y})^2 = 1020{\cdot}6$ எனக் கொடுக்கப்பட்டுள்ளது. x தனித்த மாறியாக இருக்கு மிடத்துக் கணக்கிடப்பட்ட தொடர்பு கெழு குறிப்பிடும்படியாக அமைந்துள்ளதா ?

15. 8, 7 உறுப்புகள்கொண்ட இரு தனித்த மாதிரிகளில் ஒரு குறிப்பிட்ட பண்பை அளந்ததில் பின்வரும் மதிப்புகள் கிடைத்தன.

மாதிரி 1 : 9 11 13 11 15 9 12 14

மாதிரி 2 : 10 12 10 14 9 8 10

இரண்டு மாதிரிகளும் ஒரே முழுமைத் தொகுதியைச் சேர்ந்தனவா என்பதற்கான சோதனையைச் செய்.

13. கை-வர்க்கப் பரவல்

(Chi-Square Distribution)

கருதுகோள் சோதனையில் மிக முக்கியம் வாய்ந்த மற்றொரு பரவல் கை-வர்க்கப் பரவலாகும். ஹெல்மர்ட் (Helmert, *1876*), கார்ல் பியர்ஸன் (Karl Pearson, *1900*) என்பவர்களால் தனித் தனியே விளக்கப்பட்ட இப் பரவல், இணைச் சிறப்புநிலைச் சோதனைகள் (test of goodness fit), தொடர்பற்றநிலைச் சோதனைகள் (test of independence), ஒருபடித்தான நிலைச் சோதனைகள் (test of homogenity) ஆகியவற்றின் தன்மைக்குப் பயன்படும் நிலையை இங்குக் காண உள்ளோம்.

கூட்டுச் சராசரி பூச்சியம், திட்டவிலக்கம் ஒன்று என்றுள்ள இயல்நிலை முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட x_1 x_2..........x_n எனும் சார்பற்ற (independent) மாறிகளின் கூட்டு மாறியின் ($U = x_1 + x_2 + \ldots\ldots\ldots + x_n$) பரவலும் ஓர் இயல் நிலைப் பரவலாகும். இம் மாறியின் வர்க்கங்களின் கூட்டுத் தொகை $Z = (x_1^2 + x_2^2 + \ldots\ldots\ldots x_n^2)$ ஒரு கை–வர்க்கப் (Chi-square) பரவலாக அமைகிறது.

$$y = y_0 \; e^{-\frac{1}{2} \chi^2} \; (\chi^2)^{\frac{\delta - 1}{2}}$$

δ என்பது கட்டின்மை எண்ணிக்கையைக் குறிக்கும். கை–வர்க்கப் பரவலின் வளைகோடு கட்டின்மை எண்ணிக்கைக்கேற்ற வாறு அமைப்பில் மாறுபட்டு $(0, \infty)$ இடைவெளிப் பிரிவில் வரையறுக்கப்படும். δ-வின் மதிப்பு அதிகரிக்கும்போது χ^2-ன் வளைகோடுகள் சமச்சீரான நிலைகளை அடைந்து இறுதியாக δ-வின் மதிப்பு எண்ணிலியை (∞–infinity) நெருங்க இயல்நிலை வளை கோட்டை (normal curve) நெருங்கும். χ^2-ன் மதிப்புகளில் உள்ள ஒவ்வொரு முதலடுக்குக் கட்டுமானமும் (linear constraints)

கை-வர்க்கப் பரவலின் கட்டின்மை எண்ணிக்கைக்கான அளவில் ஒன்றைக் குறைப்பதாகக் கொள்ளப்படும். x_1, x_2.........x_n என்ற இயல்நிலை மாறிகளின் வர்க்கங்களின் கூட்டுத்தொகைக்கான மாறி m முதலடுக்குக் கட்டுமானத்தால் கட்டுப்படுத்தப்படும் போது $z = x_1^2 + x_2^2 + \ldots\ldots + x_n^2$ எனும் மாறி $\gamma = (n - m)$ கட்டின்மை எண்ணிக்கைக்கு ஏற்ற கை-வர்க்கப் பரவலைப் பெற்றிருக்கும்.

கை-வர்க்கச் சோதனையை உபயோகப்படுத்தும்போது அடியிற்காணும் சில முக்கியமான கட்டளைகளை மனத்தில் கொள்ள வேண்டும்.

1. மாதிரியின் உறுப்புகள் தொடர்பற்றிருக்க வேண்டும்.
2. கட்டுமானங்கள் ஒருபடித்தானதாக இருக்கவேண்டும்.
3. அறை/பிரிவு நிகழ்வெண்கள் மிகக் குறைவாக இருக்கக் கூடாது.

கை-வர்க்கப் பட்டியல், குறிப்பிட்ட கட்டின்மை எண்ணிக்கைக்கான, நமது தேவைக்கேற்ற சிறப்பெல்லைகளில் ராண்டம் மாதிரியின் ஏற்றத்தாழ்வுகளுக்கான உச்ச மதிப்புகளைத் தருகிறது. [கை–வர்க்கப் பரவலுக்கான மாறுபட்ட கட்டின்மை எண்ணிக்கையின் ·05, ·01 சிறப்பெல்லைக்கான மதிப்புகள் பின் இணைப்பில் தரப்பட்டுள்ளன]; கணக்கிடப்படும் மதிப்பு, பட்டியல் மதிப்பை விட அதிகமாயின் கருதுகோளை மறுத்தும், குறைவாயின் கருதுகோளை ஏற்றும் முடிவு செய்யப்படும். கட்டின்மை எண்ணிக்கை முப்பதிற்கான கை-வர்க்கப் பட்டியல் எல்லாவிடத்தும் கொடுக்கப்பட்டிருப்பதைக் காணலாம். கட்டின்மை எண்ணிக்கை 30-க்கு அதிகமாக அமையும்போது $\sqrt{2\chi^2}$ எனும் மாறி, இயல்நிலைப் பரவலை ஒத்து $\sqrt{2n-1}$ என்ற கூட்டுச் சராசரியையும், ஒன்று என்ற திட்டவிலக்கத்தையும் பெறுவதால், $(\sqrt{2\chi^2} - \sqrt{2n-1})$ என்ற மாறி இயல்நிலைப் பரவலை ஒத்து அமையும். எனவே, கட்டின்மை எண்ணிக்கை 30-க்கு அதிகமாக இருக்கும்போது, இயல்நிலைப் பரவலின் நிகழ்திற மதிப்புகளோடு ஒப்பிட்டு முடிவுகள் மேற்கொள்ள ஏதுவாய் உள்ளது.

கண்டறிந்த நிகழ்வெண்களை (observed frequencies) fo_i எனவும், கோட்பாட்டு (theoritical) அல்லது கணக்கியலாக எதிர்பார்க்கும் (expected) நிகழ்வெண்களை fe_i எனவும் கொள்வோம். ஒவ்வொரு பிரிவிலும் (class) அல்லது அறையிலும் (cell) கண்

டறிந்த நிகழ்வெண்ணிற்கும் எதிர்பார்க்கப்படும் நிகழ்வெண்ணிற்கும் உள்ள வித்தியாச வர்க்கத்தை எதிர்பார்க்கப்படும் நிகழ்வெண்ணல் வகுத்துவரும் மதிப்புகளின் கூட்டுத்தொகை கை-வர்க்க மதிப்பிற்குச் சமமாகும்.

$$\chi^2\delta = \frac{(fo_1 - fe_1)^2}{fe_1} + \frac{(fo_2 - fe_2)^2}{fe_2} + \ldots\ldots + \frac{(fo_n - fe_n)^2}{fe_n}$$

$$\chi^2\delta = \sum_{1}^{n} \frac{(fo_i - fe_i)^2}{fe_i}$$

இங்கு 'δ' என்பது வரையற்ற பாகைகளைக் குறிக்கும். நிகழ்வெண்கள் $(fo_i - fe_i)$, $i = 1, 2, \ldots\ldots n$ முறையே சமமாக இருக்குமானல் χ^2-ன் மதிப்புப் பூச்சியமாகும். இப்படி அமைவது மிக அரிதாகும். இவ் விருவகை (சோதனை நிகழ்வெண், எதிர்பார்க்கும் நிகழ்வெண்) நிகழ்வெண்களுக்கான வித்தியாசத்தைச் சோதிப்பதே கை-வர்க்கச் சோதனையின் நோக்கமாகும். கை-வர்க்க வாய்பாட்டில் வித்தியாசங்களின் வர்க்கங்களை எடுத்திருப்பதால் கை-வர்க்கத்தின் மதிப்புப் பூச்சியத்திற்கு அதிகமாக அமைந்து வித்தியாசத்தின் மதிப்பு அதிகமாக இருக்கும்போது கை-வர்க்க அளவும் அதிகமாக அமையும். கொடுக்கப்பட்ட விவரங்களுக்கான கை-வர்க்க மதிப்பைக் கணக்கிட்டு, கூறெடுத்தலினல் ஏற்படும் ஏற்ற இறக்கங்களுக்கான பட்டியல் மதிப்புடன் ஒப்பிட்டுக் கருதுகோளை ஏற்றும் மறுத்தும் முடிவுசெய்கிறேம்.

காட்டு :

மக்காச்சோள இனம் ஒன்றிற்கான ஒட்டுச் சேர்க்கையில் கிடைத்த முடிவு 773 பச்சை, 231 தங்கநிறம், 238 மங்கிய பச்சை, 59 மங்கிய தங்கப் பச்சை எனக் குறிக்கப்பட்டுள்ளது. 9 : 3 : 3 : 1 என அமையும் மெண்டீலியன் முழுமைத்தொகுதியிலிருந்து இம்மாதிரி எடுக்கப்பட்டுள்ளதெனக் கொள்ள முடியுமா என்பதற்கான சோதனையை அமைப்போம்.

கொடுக்கப்பட்ட அளவுகளை fo_1, fo_2, fo_3, fo_4 எனக் கொள்வோம். கொடுக்கப்பட்டுள்ளெ மெண்டீலியன் முழுமைத் தொகுதியில் இவை அமைவதற்தான நிகழ்திறம் $\frac{9}{16}$, $\frac{3}{16}$, $\frac{3}{16}$, $\frac{1}{16}$ எனக் கொடுக்கப்பட்டுள்ளது. இந் நிகழ்திறன்களின் உதவியால் எதிர்பார்க்கப்படும் நிகழ்வெண்கள் கணக்கிடப்படுகின்றன.

fo_1	773
fo_2	231
fo_3	238
fo_4	59

fe_1	1301 × 9/16	731·9
fe_2	1301 × 3/16	243·9
fe_3	1301 × 3/16	243·9
fe_4	1301 × 1/16	81·3

சோதனைமூலம் கிடைத்த முடிவுகளுக்கும், மெண்டலியன் கருதுகோட்படி அமையும் முடிவுகளுக்கும் உள்ள வேறுபாடு குறிப்பிடும்படி உள்ளதா என்பதே நமது சோதனையாகும். இரண்டிற்குமான வேறுபாடு குறிப்பிடும்படி அமையாமல், கொடுக்கப்பட்ட மாதிரி மெண்டலியன் முழுமைத் தொகுதியைச் (9 : 3 : 3 : 1) சேர்ந்தது எனக்கொண்டு கை-வர்க்க அளவைக் கணக்கிடுகிறோம்.

fo_1	fe_1	$\frac{(fo_i - fe_i)^2}{fe_i}$	χ^2
773	731·9	$(41{\cdot}1)^2/731{\cdot}9$	2·31
231	243 9	$(12{\cdot}9)^2/243{\cdot}9$	0·68
238	243 9	$(5{\cdot}9)^2/243{\cdot}9$	0·14
59	81·9	$(22{\cdot}3)^2/81{\cdot}3$	6·12
			9·25

கணக்கிடப்பட்ட நான்கு நிகழ்வெண்களில் மூன்று நிகவெண்கள் மட்டும் தன்னிச்சைப்படி மதிப்புகளைப் பெறுவதையும், நான்காவது நிகழ்வெண் மொத்தத்திலிருந்து மற்ற மூன்று நிகழ்வெண்களின் கூட்டுத்தொகையைக் கழித்துவரும் மதிப்பிற்குச் சமமாக அமையவேண்டிய கட்டளையுடன் இருப்பதையும் உணர்வது தேவையானதாகும். இது நாம் கணக்கிட்ட கை - வர்க்கப் பரவலின் கட்டின்மை எண்ணிக்கையைத் தருகிறது. நான்கு நிகழ்வெண்களில் மூன்றுமட்டும் தன்னிச்சைப்படி அமைவதால் கை - வர்க்கப் பட்டியலில் கட்டின்மை எண்ணிக்கை மூன்றிற்கான 0·05 நிகழ்திற மதிப்பைக் குறித்துக்கொள்கிறோம்.

$\chi^2_{.05}(3) = 7{\cdot}81.$

மாதிரி எடுத்தலின் ஏற்ற இறக்கத்தின் அதிகப்படியான மூன்று வரையற்ற பாகைகளுக்கான மதிப்பு 7·81 என அமைந்து, கொடுக்கப்பட்ட மாதிரியின் அளவுகள் 9 : 3 : 3 : 1 என்ற மெண்டலியன் முழுமைத் தொகுதியிலிருந்து எடுக்கப் பட்டிருக்கும் என்ற நமது கருதுகோள் மறுக்கப்படுகிறது. மாதிரி யின் அளவுகளுக்கும் 9 : 3 : 3 : 1 எனும் மெண்டலியன் கருது கோட்படி எதிர்பார்க்கும் அளவுகட்கும் உள்ள வித்தியாசம் குறிப் பிடும்படி அமைந்துள்ளது. ஏற்பட்ட வித்தியாசம் மாதிரி எடுத்த லின் ஏற்ற இறக்கத்திற்கானதாகக் கொள்ள முடியாதாகையால், கொடுக்கப்பட்ட மாதிரி 9 : 3 : 3 : 1 எனும் மெண்டலியன் முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதெனக் கொள்ள இடமளிக்கவில்லை என்ற முடிவைச் சொல்கிறோம்.

காட்டு :

120, 48, 36, 13 என அமையும் மற்றொரு மாதிரியின் முடிவு கள் 9 : 3 : 3 : 1 எனும் மெண்டலியன் முழுமைத்தொகுதியைச் சேர்ந்ததா என்பதற்கான சோதனையைச் செய்வோம்.

fo	fe_i	$\frac{(fo_i - fe_i)^2}{fe_i}$
$fo_1 = 120$	$fe_1 = 217 \times \frac{9}{16} = 122\cdot06$	$(-2\cdot06)^2/122\cdot06 = 0\cdot0347$
$fo_2 = 48$	$fe_2 = 217 \times \frac{3}{16} = 40\cdot69$	$(-7\cdot31)^2/40\ 69 = 1\cdot3130$
$fo_8 = 36$	$fe_3 = 217 \times \frac{3}{16} = 40\cdot69$	$(-4\ 69)^2/40\cdot69 = 0.5405$
$fo_4 = 13$	$fe_4 = 217 \times \frac{1}{16} = 13\cdot56$	$(-0\cdot56)^2/13\cdot56 = 0\cdot0231$
217		1·9113

கட்டின்மை எண்ணிக்கை மூன்றிற்கான 0·05 நிகழ்திறக் கை-வர்க்க மதிப்பு $\chi^2_{\cdot05(3)} = 7\cdot81$. கணக்கிடப்பட்ட மதிப்பு, பட்டியல் மதிப்பைவிடக் குறைவாக இருப்பதால் மாதிரியின் அளவுகளுக்கும் 9 : 3 : 3 : 1 எனும் மெண்டலியன் கருதுகோட்படி

எதிர்பார்க்கப்படும் அளவுகளுக்கும் உள்ள வேறுபாடு குறிப்பிடும் படி அமையாமல், அமைந்த வித்தியாசம் மாதிரி எடுத்தலின் ஏற்றத் தாழ்வினல் ஏற்பட்டதெனக்கொண்டு, கொடுக்கப்பட்ட மாதிரி 9 : 3 : 3 : 1 எனும் மெண்டலியன் முழுமைத்தொகுதியைச் சேர்ந்ததெனக்கொள்ள இடமளிக்கின்றது என்ற முடிவைச் சொல்கின்ருேம்.

இணைச் சிறப்புநிலைச் சோதனை (Test of goodness of fit)

N உறுப்புகள் கொண்ட முழுமைத் தொகுதியை k இனங்களாகப் பிரிப்பதில் ஒவ்வோர் உறுப்பும் ஏதாவது ஒரு பிரிவில் ஒரு தடவைமட்டும் அமையப்பெறும் எனக் கொள்வோம். ஒவ்வொரு பிரிவு நிகழ்வெண்களையும் $fo_1, fo_2, \ldots\ldots fo_k$ எனக் கொள்வோம். $\sum_{i=1}^{k} foi = N$ என அமையும்படி நமது பிரிவுகள் அமைந்துள்ளதைக் காணலாம். இம் முறைப்படி அமைந்த பிரிவெண்களுக்கான கருதுகோள் நிகழ்வெண்கள் தரும் பரவல் களை அறிவதில் நமது நாட்டம் செல்கிறது. கருதுகோள் நிகழ்

பிரிவு	சோதனை நிகழ்வெண் fo_i	கருதுகோள் (எதிர்பார்க்கப்படும்) நிகழ்வெண் fe_i
1	fo_1	fe_1
2	fo_2	fe_2
3	fo_3	fe_3
.	.	.
.	.	.
.	.	.
k	fo_k	fe_k
	N	N

வெண்கள் $fe_1, fe_2, \ldots\ldots fe_k$ என அமைந்து $\sum_{i=1}^{k} f_{ci} = N$ என்ற கட்டளைக்குட்படும்போது, இவ்விரு நிகழ்வெண் பகுதிகளும் ஒத்ததாய் அமைகிறதா என்பதே நமது சோதனையாகும்.

$$\text{சோதனை அளவு}: \quad \chi^2 = \sum_{i=1}^{k} \frac{(fo - fe_i)^2}{fe_i}$$

இக் கை-வர்க்கப் பரவல் $(k - 1)$ வரையற்ற பாகைகளுக்கானதாகக் கொள்ளப்படும். எதிர்பார்க்கப்படும் நிகழ்வெண்களின் கூட்டுத்தொகை சோதனை நிகழ்வெண்களின் கூட்டுத்தொகைக்குச் சமமாக அமையும் நியதி, k பிரிவுகளில் அமையும் எதிர்பார்க்கப்படும் நிகழ்வெண்களில் $(k - 1)$ பிரிவுகளில்தான் தன்னிச்சைப்படி அமையும். இறுதியாக (எந்தப் பிரிவாயினும்) அமையும் நிகழ்வெண்களின் மதிப்பு, மற்ற மதிப்புகளைச் சார்ந்தமைந்து கை - வர்க்கப் பரவலுக்கான k-மதிப்புகளில் கட்டின்மை எண்ணிக்கையில் ஒன்றைக் குறைக்கிறது. கட்டின்மை எண்ணிக்கை மொத்த உறுப்புகளான N ஐச் சாராமல் அவை பிரிக்கப்படும் k பிரிவுகளைச் சார்ந்துள்ளதை உணர்வது அவசியம். எதிர்பார்க்கப்படும் நிகழ்வெண்களைக் கணக்கிடும் போது கருதுகோள் பரவலுக்கான அளவைகள் கணக்கிடப் பட்டால், கணக்கிடப்படும் ஒவ்வோர் அளவைக்கும் (statistic) கட்டின்மை எண்ணிக்கையில் ஒன்று குறைக்கப்பட வேண்டியதை மனத்தில் கொள்ளவேண்டும்.

கொடுக்கப்பட்ட நிகழ்வெண் பரவலுக்கான மாறி ஈருறுப்புப் பரவலைத் தழுவியதாகக் கொண்டு 'p'-மதிப்பிடப்பட்டால், கணக்கிடப்படும் கை - வர்க்கப் பரவல் $(k - 1 - 1) = (k - 2)$ கட்டின்மை எண்ணிக்கையைப் பெற்றதாய் அமையும். எதிர் பார்க்கப்படும் நிகழ்வெண்களின் கூட்டுத்தொகை கொடுக்கப் பட்ட நிகழ்வெண்களின் கூட்டுத்தொகைக்குச் சமமாக இருக்க வேண்டும் என்ற நிபந்தனைக்கொன்றும், 'p' மதிப்பிடப்பட்டதற்கொன்றுமாக இரண்டு கட்டின்மை எண்ணிக்கை k-லிருந்து கழிக்கப்பட்டு, கணக்கிடப்பட்ட கை-வர்க்கப் பரவல் $(k - 2)$ கட்டின்மை எண்ணிக்கை கொண்டதாகக் கருதப்படும். நிகழ்வெண் பரவலின் மாறி இயல்நிலைப் பரவலைத் தழுவியதாகக் கொண்டு கூட்டுச் சராசரி $(\bar{x})$, திட்டவிலக்கம் (σx) இரண்டும் கணக்கிடப்பட்டபின் கணக்கிடப்படும் கை-வர்க்கம் $(k - 1 - 2) = (k - 3)$ கட்டின்மை எண்ணிக்கையைக்கொண்டு அமையும்.

காட்டு :

பாய்சான் பரவலைக் கருதுகோளாகக் கொண்டு கணக்கிடப்பட்ட எதிர்பார்க்கப்படும் நிகழ்வெண்கள் கொடுக்கப்பட்ட (ஈருறுப்பு மற்றும் பாய்ஸான் பரவல்) மாதிரிக்கு இணைச் சிறப்புநிலைச் சோதனையை அமைப்போம்.

x	fo	fei	$(fo_i - fe_i)^2$	$\frac{(fo_i - fe_i)^2}{fe_i}$
0	103	107	16	0·1495
1	143	141	4	0·0284
2	98	93	25	0·2688
3	42	41	1	0·0244
4	8	14	36	2·5714
5	4 }	4 }	1	0·2000
6	2 }	1 }	.	.
7	.	.	.	.
8	.	.	.	.
9	.	.	.	.
10	.	.	.	.
				3·2425

கருதுகோள் நிகழ்வெண்களைப் பாய்ஸான் நியதிப்படி அமைக்கும் முறையை முன்பு கண்டோம். நமது கருதுகோள் உண்மையாக இருக்குமாயின், கொடுக்கப்பட்ட நிகழ்வெண்களும் எதிர்பார்க்கப்படும் நிகழ்வெண்களும் சமமாக அமைந்து, கை-வர்க்கப் பரவலின் மதிப்புப் பூச்சியமாக அமைந்திருக்கும். நடைமுறையில் இப்படி அமைவது மிக அரிதாகும். எதிர்பார்க்கப்படும் நிகழ்வெண்களுக்கும், கொடுக்கப்பட்ட நிகழ்வெண்களுக்கும் உள்ள வித்தியாசம் குறிப்பிடும்படி அமையாமல், இப் பரவல் பாய்ஸான் நியதியைச் சார்ந்துள்ளதா அல்லது வித்தியாசம் குறிப்

பிடும்படி அமைந்து பாய்சான் நியதியைச் சார்ந்ததாகக்கொள்ள முடியாதா என்பதே நமது சோதனையாகும். இப் பரவலை அமைக்கும்போது இதற்குத் தேவையான 'm'-ன் மதிப்பைக் கணக்கிட்டு எதிர்பார்க்கப்படும் நிகழ்வெண்களைக் கணக்கிட்டோம். x மாறி 5, 6 என்ற மதிப்புகளைப் பெறும்போது (4,2), (4,1) எனக் கொடுக்கப்பட்ட எதிர்பார்க்கப்படும் நிகழ்வெண்கள் பெறுவதை அறிகிறோம். நிகழ்வெண்கள் குறைந்தது ஐந்து இருக்கவேண்டுமென்ற கை-வர்க்கப் பரவலின் சிறப்பு நியதிக்கான தன்மைக்கு, இவ்விரு நிகழ்வெண்களையும் கூட்டிக் கை-வர்க்க மதிப்புக் கணக்கிடப்படுகிறது. கூட்டி அமைக்கப்பட்டபின் நிகழ்வெண்கள் 6 (k) பிரிவுகளில் அமைவதை உணர்கிறோம். கணக்கிடப்பட்ட கை-வர்க்க மதிப்பு 3·2425. ஒரு பிரிவு தன்னிச்சைப்படி அமைய முடியாத தன்மைக்கு ஒரு கட்டின்மை எண்ணிக்கையும், இந் நிகழ்வெண்கள் கணக்கிட மதிப்பிட m-ன் அளவைக்காக ஒரு கட்டின்மை எண்ணிக்கையும் இழந்து, மதிப்பிட்ட கை-வர்க்க அளவு $4 = (6-1-1)$ கட்டின்மை எண்ணிக்கையுடன் ஒரு கை-வர்க்கப் பரவலாக அமைகிறது. $\chi^2_{\cdot 05\,(4)} = 9{\cdot}488$. கொடுக்கப்பட்ட எதிர்பார்க்கப்படும் நிகழ்வெண்களுக்கான வித்தியாசம் குறிப்பிடும்படி அமையாமல், இவை பாய்ஸான் நியதியைச் சார்ந்துள்ளதெனக் கொள்ள இடமிருக்கிறது என்ற முடிவைச் சொல்கிறோம்.

காட்டு :

'ஈருறுப்பு மற்றும் பாய்ஸான்' பிரிவில் பொருத்தப்பட்ட ஈருறுப்புப் பரவலுக்கான நியதி பொருத்தமானதா என்ற சோதனையைப் (317 ஆம் பக்கம்) பார்ப்போம்.

ஈருறுப்புப் பரவலைப் பொருத்துவதற்காகக் கணக்கிடப்பட்ட $\bar{x} = np$, மற்றும் நிகழ்வெண்களின் கூட்டுத்தொகை சமமாக அமைவதற்கான கட்டளை ஆகிய இரண்டிற்கும் இரு கட்டின்மை எண்ணிக்கையைக் குறைத்து $\chi^2_{\cdot 05(4)} = 9{\cdot}488$ என்ற பட்டியல் மதிப்புடன் கணக்கிடப்பட்ட மதிப்பை ஒப்பிட்டு முடிவு செய்கிறோம்.

கருதுகோள் சோதனையில் அமையும் நிகழ்வெண்களுக்கும் கொடுக்கப்பட்ட நிகழ்வெண்களுக்கும் உள்ள வித்தியாசம் குறிப்பிடும்படி அமையாமல், ஈருறுப்பு நியதிப்படி அமைகிறது எனும் நமது கருதுகோள் ஏற்கப்பட்டு, கொடுக்கப்பட்ட பரவல் ஓர் ஈருறுப்புப் பரவலை ஒத்து அமைகிறது என்னும் நியதியை ஏற்றுக்கொள்கிறோம்.

x_i	f_o	f_{ei}	$(f_{oi}-f_{ei})$	$(f_o-f_{ei})^2$	$(f_{oi}-f_{ei})^2/f_e$
0	6	6·9	0·9	0·81	0·1174
1	20	19·1	0·9	0·81	0·0424
2	28	24·0	4·0	16·00	0·5714
3	12	17·8	5·8	33·64	1·8899
4	8	8·6	0·6	0·36	0·0419
5	6	2·9 }	2·3	5·29	1·4297
6	0	0·7 }	.	.	.
7	0	0·1 }	.	.	.
8	0	.	.	.	.
9	0	.	.	.	.
10	0	.	.	.	.
	80	80·1	.	.	4·0927

2 × 2 இணைப்புப் பட்டியல் (2 × 2 Contingency table)

சோதனை முடிவுகளுக்கும், நாம் எதிர்பார்க்கும் அளவுகளுக்கும் ஒற்றுமை (agreement) இருக்கிறதா எனும் சோதனைகள் சார்பற்ற தன்மைக்கான சோதனை முறைகளைச் சாரும். ஒரு முழுமைத் தொகுதியைச் சேர்ந்த உறுப்புகள் இரண்டு முறைகளில் இரு கூறுகளாகப் பிரிக்கப்படும்போது, இவ் விரண்டு கூறுகளும் தொடர்புடையனவா எனும் சோதனை சார்பற்ற (independent) தன்மைக்கான சோதனையாகும். ஒரு முழுமைத் தொகுதியை, ஒரு குறிப்பிட்ட வியாதிக்கான தடுப்பூசி போட்டவர்கள் தடுப்பூசி போடாதவர்கள் என்றும், தடுப்பூசி போட்டும் வியாதியால் பாதிக்கப்பட்டவர்கள், வியாதியால் பாதிக்கப்படாதவர்கள் என்றும் இருவகையாகப் பிரித்துள்ளதாகக் கொள்வோம். இவ்வாறு பிரித்தமைக்கப்படும் பட்டியல் 2 × 2 இணைப்புப் பட்டியல் எனப்படும்.

	வியாதியால் பாதிக்கப் பட்டவர்	வியாதியால் பாதிக்கப் படாதவர்	
தடுப்பூசி போட்டுக் கொண்டவர்கள்	a	c	$a+c$
தடுப்பூசி போடா தவர்கள்	b	d	$b+d$
	$a+b$	$c+d$	$a+b+c+d$

'a' என்பது தடுப்பூசி போட்டும் வியாதியால் பாதிக்கப் பட்டவர்களேயும், 'b' என்பது தடுப்பூசி போடாமல் வியாதியால் பாதிக்கப்பட்டவர்களேயும், 'c' என்பது தடுப்பூசி போட்டு வியாதியால் பாதிக்கப்படாதவர்களேயும், 'd' என்பது தடுப்பூசி போடாமல் வியாதியால் பாதிக்கப்படாதவர்களேயும் குறிக் கின்றன. நமது சோதனே தடுப்பூசி போடுவதற்கும், வியாதி யால் பாதிக்கப்படாமலிருப்பதற்கும் தொடர்புண்டா என் பதைச் சோதிப்பதாகும்.

'தடுப்பூசி போடுவதற்கும் வியாதியால் பாதிக்கப்படா மலிருப்பதற்கும் தொடர்பில்லே.' அதாவது தடுப்பூசியும் வியாதியும் தொடர்பற்றவை எனும் கருதுகோளே அமைத்து, கருதுகோளே ஏற்ருே மறுத்தோ முடிவு செய்ய உள்ளோம்.

ஒவ்வோர் அறைக்கும் நாம் எதிர்பார்க்கும் நிகழ்வெண்கள் கணக்கிடப்பட்டு, கை-வர்க்க அளவை மதிப்பிடப்படுகிறது. பட்டியல் மதிப்போடு இதனே ஒப்பிட்டு முடிவு செய்கின்ருேம், சோதனே முடிவும் நாம் எதிர்பார்க்கும் அளவுகளின் மொத்தமும் சமமாக இருக்கும் கட்டளேப்படி 2×2 இணேப்புப் பட்டியலுக்கான நான்கு அறைகளில் ஏதேனும் ஒன்றிற்கான எதிர்பார்க்கும் எண்ணே அறிந்து, துணே மொத்தங்களின் (sub-totals) உதவியால் மற்ற அறை எண்கள் கணக்கிடப்படுகின்றன. 2×2 இணேப்புப் பட்டியலில் ஓர் அறையெண் மட்டுமே தன்னிச்சைப்படி அமை வதால், இதற்கான கை-வர்க்கப் பரவலின் கட்டின்மை எண் ணிக்கை ஒன்ருக அமைகிறது. கணக்கிடப்படும் கை-வர்க்க மதிப்பை ·05 சிறப்பெல்லேயில் கட்டின்மை எண்ணிக்கை ஒன்றிற்கான கை-வர்க்கப் பட்டியல் மதிப்புடன் ஒப்பிட்டுக் கருது கோளே ஏற்றும் மறுத்தும் முடிவு செய்கின்ருேம்.

காட்டு :

343 நபர்களின் 2×2 இணைப்புப் பட்டியல் தரப்பட்டுள்ள சார்பற்ற தன்மைக்கான சோதனையைக் காண்போம்.

	வியாதியால் பாதிக்கப் பட்டவர்கள்	வியாதியால் பாதிக்கப் படாதவர்கள்	மொத்தம்
தடுப்பூசி போட்டவர்கள்	192	4	196
தடுப்பூசி போடாதவர்கள்	113	34	147
	305	38	343

ஓர் அறையில் அமையும் எதிர்பார்க்கப்படும் நிகழ்வெண்ணைக் கணக்கிட்டு, துணை மொத்தங்களிலிருந்து இவ்வெண்ணைக் கழித்து மற்ற அறைகளுக்கான நிகழ்வெண்களைக் கணக்கிடுகிறோம்.

முதல் அறையில் நாம்

எதிர்பார்க்கும் நிகழ்வெண் $= \frac{305 \times 196}{343} = 174{\cdot}28$

இரண்டாம் அறையில் நாம்

எதிர்பார்க்கும நிகழ்வெண் $= 305 - 174{\cdot}28$

$= 130{\cdot}72$

மூன்றும் அறைக்கான நிகழ்வெண் $= 196{\cdot}0 - 174{\cdot}28 = 21{\cdot}72$

நான்காம் அறைக்கான நிகழ்வெண் $= 147{\cdot}0 - 130{\cdot}7 = 16{\cdot}28$

எதிர்பார்க்கும் நிகழ்வெண்களுக்கான பட்டியல்

	வியாதியால் பாதிக்கப் பட்டவர்கள்	வியாதியால் பாதிக்கப் படாதவர்கள்	
தடுப்பூசி போட்டவர்கள்	174·28	21·72	196
தடுப்பூசி போடாதவர்கள்	130·72	16·28	147
	305	·38	343

$$\chi^2 = \frac{(192 - 174{\cdot}28)^2}{174{\cdot}28} + \frac{(113 - 130{\cdot}72)^2}{130{\cdot}72}$$

$$+ \frac{(4 - 21{\cdot}72)^2}{21{\cdot}72} + \frac{(34 - 16{\cdot}28)^2}{16{\cdot}28}$$

$$= 1{\cdot}80 + 2{\cdot}40 + 14{\cdot}46 + 19{\cdot}29$$

$$= 37{\cdot}91$$

கணக்கிடப்பட்ட கை-வர்க்க மதிப்பு, கட்டின்மை எண்ணிக்கை ஒன்றிற்கான கை-வர்க்கப்பட்டியல் மதிப்பைவிட அதிகமாக இருப்பதால் தடுப்பூசி போடுவதற்கும் வியாதியால் பாதிக்கப்படுவதற்கும் தொடர்பில்லை என்னும் கருதுகோள் மறுக்கப்பட்டு இரண்டிற்கும் தொடர்புண்டு எனும் நிலையைச் சொல்கிறோம்.

எதிர்பார்க்கப்படும் நிகழ்வெண்கள் கணக்கிடப்படாமலேயே 2×2 இணைப்புப் பட்டியலுக்கான கை-வர்க்க அளவு

$$\chi^2 = \frac{(ad - bc)^2\,(a + b + c + d)}{(a + b)\,(c + d)\,(a + c)\,(b + d)}$$

என்ற வாய்பாட்டிலிருந்து கணக்கிட்டுக் கொள்வதும் நடைமுறையில் உள்ளது. கொடுக்கப்பட்ட கணக்கிற்கு இம் முறைப்படி கணக்கிடும்போது

$$\chi^2 = \frac{(192 \times 34 - 113 \times 4)^2\,(343)}{(305)\,(38)\,(196)\,(147)} = 37{\cdot}91$$

இரண்டு முறைகளிலும் ஒரே மாதிரியான முடிவுகளே எடுக்க நேரிடுவதால், கை-வர்க்க மதிப்பை மேற்கண்ட வாய்பாட்டின் மூலம் கணக்கிடுவதால், எதிர்பார்க்கும் நிகழ்வெண்கள் கணக்கிடுவதைத் தவிர்க்கலாம்.

யேட்சின் திருத்தம் (Yates Correction)

ஒரு தொடர்மாறிக்கான மதிப்புகளையே கை-வர்க்கப் பட்டியல் தருகிறது. நிகழ்வெண் பரவலும் தொடர்ச்சியாக இருந்தால்தான், இப் பட்டியல் மதிப்போடு ஒப்பிடும்போது சிறந்த

முடிவுகளை எடுக்க முடியும். அறையெண்கள் மிகவும் குறைவாய் அமையும்போது ஏற்படும் தவற்றினைச் சரிசெய்ய 2×2 இணைப்புப் பட்டியலில் ஒரு திருத்தம் செய்கிறோம். ஐந்திற்குக் குறைவாக அமைந்த நிகழ்வெண்களை, மற்றப் பிரிவுகளுடன் சேர்த்து, கை-வர்க்க மதிப்புக் கணக்கிடப்பட்டதை அறிவோம். 2×2 இணைப்புப் பட்டியலில் இப்படிச் சேர்க்கும் முறை முடியாதாகையால், இதற்கான திருத்தத்தை ஏற்படுத்த வேண்டியவர்களாக உள்ளோம். 2×2 இணைப்புப் பட்டியலுக்கான யேட்சின் திருத்தத்தை (Yates correction) இங்குக் காண உள்ளோம்.

a, b, c, d எனும் நான்கறை எண்களின் மதிப்பைக்கொண்டு ad, bc மதிப்புகளைக் கணக்கிட்டுக் கொள்கிறோம். அதிகமான மதிப்பைத் தரும் இரு எண்களிலிருந்து $\frac{1}{2}$ ஐக் கழித்தும், குறைவான மதிப்பைத் தரும் இரு எண்களுடன் $\frac{1}{2}$ ஐக் கூட்டியும் இணைப்புப் பட்டியலை மாற்றி அமைக்கிறோம். மாற்றி அமைக்கப்பட்ட பட்டியலின் துணை மொத்தங்களும் மொத்தமும் கொடுக்கப்பட்ட பட்டியலுக்கு ஒத்து அமையவேண்டியது குறிப்பிடத்தக்கது. இவ்வாறு திருத்தி அமைக்கப்பட்ட பட்டியலுக்கான கை-வர்க்க அளவே யேட்சின் திருத்தப்பட்ட கை-வர்க்க மதிப்பு எனப்படும்.

காட்டு :

	A	B	
I	10 a	4 c	14 $a+c$
II	2 b	8 d	10 $b+d$
	12 $a+b$	12 $c+d$	24 $a+b+c+d$

$$\chi^2 = \frac{(ad - bc)^2 (a + b + c + d)}{(a + b)(c + d)(a+c)(b + d)}$$

$$= \frac{24\,(80 - 8)^2}{12 \times 12 \times 14 \times 10} = 6{\cdot}17$$

இரண்டு அறைகளுக்கான நிகழ்வெண்கள் மிகக் குறைவாக இருப்பதால், அதிகமான மதிப்பைத் தரும் a.d-லிருந்து $\frac{1}{2}$ ஐக் குறைத்தும், குறைவான மதிப்பைத் தரும் b.c-யுடன் $\frac{1}{2}$ ஐக் கூட்டியும் பட்டியலை மாற்றி அமைத்துக் கை-வர்க்க மதிப்பைக் கணக்கிடுகிறோம்.

	A	B	
I	9·5	4·5	14
II	2·5	7·5	10
	12	12	24

$$\chi^2 = \frac{(71{\cdot}25 - 11{\cdot}25)^2 \times 24}{12 \times 12 \times 14 \times 10}$$
$$= 4{\cdot}29$$

திருத்தப்படுவதற்கு முன் $(ad - bc)^2 = 72^2 = 5184$ எனவும், திருத்தப்பட்ட பின்பு $(ad - bc)^2 = 60^2 = 3600$ எனவும் அமைந்து கை-வர்க்க மதிப்பில் மாற்றம் ஏற்பட்டுள்ளதை உணர்கிறோம்.

$R \times C$ இணைப்புப் பட்டியல் ($R \times C$ Contingency table)

முழுமைத் தொகுதியை 2 × 2 வழிகளில் பிரித்தமைத்த இணைப்புப் பட்டியலுக்கான சோதனையைக் கண்டோம். 2 × 2 இணைப்புப் பட்டியல் $R \times C$ என்ற பொது இணைப்புப் பட்டியலின் ஒரு குறிப்பிட்ட அமைப்பைக் குறிக்கும். முழுமைத் தொகுதியிலுள்ள உறுப்புகள் ஒரு பண்பின் அடிப்படையில் 'R' பிரிவுகளாகவும், இரண்டாவது பண்பின் அடிப்படையில் 'C' பிரிவுகளாகவும் பிரிக்கப்பட்டு அமையும் பட்டியலே $R \times C$ இணைப்புப் பட்டியல் எனப்படும். $R \times C$ மொத்த அறைகள்கொண்ட இப் பட்டியலில் உள்ள அறை நிகழ்வெண்கள் யாவும் அவ்வறையில் அமையப்பெற்ற உறுப்புகளின் எண்ணிக்கையைச் சுட்டும். துணை மொத்தங்களும் மொத்தமும் கொடுக்கப்பட்ட மற்றும் எதிர்பார்க்கப்படும் எண்களுக்குச் சமமாக அமையும்

நிபந்தனையில் ஒவ்வொரு நிமிர் நிலை வரிசையும், படுகைநிலை வரிசையும் ஒவ்வோர் அறையெண்ணைத் தன்னிச்சைப்படி அமைக்கும் நிலையைப் பெறாமல் இருக்கும். R-படுகை நிலை வரிசைகளில் $(R-1)$ வரிசைகளே தன்னிச்சைப்படி நிகழ்வெண்களைப் பெற்றும், C-நிமிர் நிலை வரிசைகளில் $(C-1)$ வரிசைகளே தன்னிச்சைப்படி நிகழ்வெண்களைப் பெற்றும் அமைவதால் $R \times C$ இணைப்புப்பட்டியலுக்கான கை-வர்க்கப் பரவல் $(R-1)(C-1)$ கட்டின்மை எண்ணிக்கையைக் கொண்டதாய் அமையும்.

காட்டு :

விவசாயப் பண்ணைகளைச் சொந்தம், குத்தகை, இரண்டும் கலந்தது என்ற அடிப்படையிலும், நிலத்தின் தரத்தை I, II, III என்ற மூன்று பிரிவுகளாகவும் பிரித்து 3×3 $(R=3,\ C=3)$ இணைப்புப் பட்டியல் அமைக்கப்படுகிறது.

	உழவர்களின் தன்மை			
நிலத்தின் தரம்	சொந்தம்	குத்தகை	இரண்டும் கலந்தது	மொத்தம்
I	36	67	49	152
II	31	60	49	140
III	58	87	80	225
மொத்தம்	125	214	178	517

நிலங்களின் தரத்திற்கும், அவற்றை உழும் உழவர்களின் தன்மைக்கும் தொடர்புண்டா என்பதே நமது சோதனையாக அமைகிறது. இவையிரண்டிற்கும் தொடர்பில்லை என்ற கருதுகோளை அமைத்துக் கை-வர்க்க மதிப்பைக் கணக்கிடுகிறோம்.

2×2 இணைப்புப் பட்டியலில் கண்டதுபோல், துணை மொத்தங்களைக் கொண்டு எதிர்பார்க்கப்படும் நிகழ்வெண்கள் கணக்கிடப்படுகின்றன. ஓர் அறையின் எதிர்பார்க்கப்படும் நிகழ்

வெண், அவ்வறை அமைந்த நிமிர்நிலை வரிசை மொத்தத்தையும், படுகைநிலை வரிசை மொத்தத்தையும் பெருக்கிவரும் தொகையை மொத்த உறுப்புகளால் வகுத்துவரும் ஈவினுக்குச் சமமாகும்.

ஓர் அறையின் எதிர்பார்க்கப்படும் நிகழ்வெண் =

$$\frac{\text{அவ்வறை அமைந்த படுகைநிலை வரிசை மொத்தம் (row total)} \times \text{அவ்வறை அமைந்த நிமிர்நிலை வரிசை மொத்தம் (column total)}}{\text{மொத்தம் (total)}}$$

முதல் படுகைநிலை வரிசைக்கான (first row) எதிர்பார்க்கப்படும் நிகழ்வெண்கள்

$$O_1 = \frac{152 \times 125}{517} = 36{\cdot}75$$

$$O_2 = \frac{152 \times 214}{517} = 62{\cdot}92$$

$$O_3 = \frac{152 \times 178}{517} = 52{\cdot}33$$

இரண்டாவது படுகைநிலை வரிசைக்கான (second row) எதிர்பார்க்கப்படும் நிகழ்வெண்

$$O_4 = \frac{140 \times 125}{517} = 33{\cdot}85$$

$$O_5 = \frac{140 \times 214}{517} = 57{\cdot}95$$

$$O_6 = \frac{140 \times 178}{517} = 48{\cdot}20$$

மூன்றுவது படுகைநிலை வரிசைக்கான (third row) எதிர்பார்க்கப்படும் நிகழ்வெண்

$$O_7 = \frac{225 \times 125}{517} = 54{\cdot}40 = (152 - 36{\cdot}75 - 33{\cdot}85)$$

$$O_8 = \frac{225 \times 214}{517} = 93{\cdot}13 = (140 - 62{\cdot}92 - 57{\cdot}95)$$

$$O_9 = \frac{225 \times 178}{517} = 77{\cdot}47 = (225 - 52{\cdot}33 - 48{\cdot}20)$$

எதிர்பார்க்கப்படும் நிகழ்வெண்களுக்கான பட்டியல்

	மொத்தம்	குத்தகை	இரண்டும் கலந்தது	மொத்தம்
I	36·75	62·92	52·33	152
II	33·85	57·95	48·20	140
III	54·40	93·13	77·47	225
மொத்தம்	125	214	178	517

$$\chi^2 = \frac{(36 - 36{\cdot}75)^2}{36{\cdot}75} + \frac{(67 - 62{\cdot}92)^2}{62{\cdot}92} + \ldots\ldots$$

$$+ \frac{(80 - 77{\cdot}41)^2}{77{\cdot}41}$$

$$= 1{\cdot}54$$

துணைமொத்தங்கள் (sub totals) கொடுக்கப்பட்ட முடிவிலும், எதிர்பார்க்கப்படும் முடிவிலும் சமமாக இருக்கவேண்டுமென்ற நிபந்தனை கட்டின்மை எண்ணிக்கையைக் கணக்கிட உதவியாக இருக்கிறது. படுகைநிலை வரிசை மூன்றிலும் உள்ள ஒன்பது அறைகளில் ஆறு அறைகள் மட்டும் தன்னிச்சைப்படி அமைந்து, மூன்று அறைகளுக்கான நிகழ்வெண்கள் துணை மொத்தங்களிலிருந்து மற்ற இரண்டறைகளின் நிகழ்வெண்களைக் கழித்துவரும் மதிப்பிற்குச் சமமாகின்றன. நிமிர்நிலை வரிசையிலும் இதுபோலவே அமைந்துள்ளது. $(R - 1)(C - 1)$ என்ற நியதிப்படி $R = 3, C = 3$ எனும் 3×3 இணைப்புப்பட்டியல் $(3 - 1)(3 - 1) = (2 \times 2) = 4$ கட்டின்மை எண்ணிக்கையைக் கொண்ட கை-வர்க்கப் பரவலின் 0·05 சிறப்பெல்லை மதிப்புடன் ஒப்பிட்டு முடிவுசெய்கிறோம்.

$$\chi^2_{\cdot 05(4)} = 9{\cdot}49$$

கணக்கிடப்பட்ட மதிப்பு (1·54), நாம் எதிர்பார்க்கும் மதிப்பை (9·49) விடக் குறைவாக இருப்பதால், நிலத்தின் தரத்திற்கும் அதனை உழுபவர்கள் தன்மைக்கும் தொடர்பில்லை எனும் கருதுகோளை ஏற்றுக்கொள்கிறோம்.

காட்டு :

கூலவகைகளில் கதிரலகுடைய (awned), கூலவகைகளில் கதிரலகற்ற (awnless) புல்லரிசிப் பயிர் (oats) வகைக்கான விளைபலன் தரப்பட்டுள்ளது.

	விளைபலன் (எடை கிராமில்)			
	151–200	201–250	251–325	மொத்தம்
கூலவகைகளில் கதிரலகுடையவை	6	7	21	34
கூலவகைகளில் கதிரலகற்றவை	18	21	9	48
	24	28	30	82

கதிரலகிற்கும், விளைபலனுக்கும் தொடர்புண்டா என்பதைச் சோதனை செய்யவேண்டியவர்களாக உள்ளோம். கதிரலகிற்கும் விளைபலனுக்கும் தொடர்பில்லை எனும் கருதுகோளை அமைத்து கை-வர்க்க மதிப்பைக் கணக்கிடுகிறோம். இதனைப் பட்டியல் மதிப்போடு ஒப்பிட்டுக் கருதுகோளை ஏற்றோ மறுத்தோ முடிவு செய்கிறோம். இப் பட்டியல் 2 × 3 எனும் முறைப்படி அமைந்துள்ளதால், துணைமொத்தங்கள் சமமாக அமையும் கட்டளைப்படி இரண்டு அறை எண்களைக் கணக்கிட்டு மற்ற அறை எண்கள் அமைக்கப்படுகின்றன.

முதல் அறைக்கான எதிர்பார்க்கப் படும் நிகழ்வெண் $\left.\right\} = \dfrac{24 \times 34}{82} = 9{\cdot}95$

இரண்டாவது அறைக்கான எதிர் பார்க்கப்படும் நிகழ்வெண் $\left.\right\} = 24 - 9{\cdot}95 = 14{\cdot}05$

மூன்றாவது அறைக்கான எதிர் பார்க்கப்படும் நிகழ்வெண் $\left.\right\} = \dfrac{28 \times 34}{82} = 11{\cdot}61$

நான்காவது அறைக்கான எதிர் பார்க்கப்படும் நிகழ்வெண் $\left.\right\} = 28 - 11{\cdot}61 = 16{\cdot}39$

ஐந்தாவது அறைக்கான எதிர்பார்க்கப்படும் நிகழ்வெண் $= 34 - 9{\cdot}95 - 11{\cdot}61$

$= 12{\cdot}44$

ஆறுவது அறைக்கான எதிர்பார்க்கப்படும் நிகழ்வெண் $= 48 - 14{\cdot}05 - 16{\cdot}39$

$= 17{\cdot}56$

எதிர்பார்க்கப்படும் நிகழ்வெண்களுக்கான பட்டியல்

	விளைபலன் (எடை கிராமில்)			
	151–200	201–250	251–325	
கதிரலகுடையவை	9·95	11·61	12·44	34
கதிரலகற்றவை	14·05	16·39	17·56	48
	24	28	30	82

$$\chi^2 = \frac{(6 - 9{\cdot}95)^2}{9{\cdot}95} + \frac{(7 - 11{\cdot}61)^2}{11{\cdot}61} + \ldots\ldots\ldots\ldots + \frac{(9 - 17{\cdot}56)^2}{17{\cdot}56}$$

$$= 1{\cdot}57 + 1{\cdot}83 + 5{\cdot}89 + 1{\cdot}10 + 1{\cdot}30 + 4{\cdot}17 = 15{\cdot}86$$

$R = 2,\ C = 3$ என்ற அமைப்புள்ள இப் பட்டியலுக்கான கை-வர்க்கப் பரவலின் கட்டின்மை எண்ணிக்கை இரண்டு $(R - 1)(C - 1) = (2 - 1)(3 - 1) = 1 \times 2 = 2$ என அமைகிறது. இரண்டறைகளுக்கான எதிர்பார்க்கப்படும் நிகழ்வெண்கள் மட்டும் கணக்கிடப்பட்டு, மற்றவை துணைமொத்தங்களிலிருந்து கணக்கிடப்பட்டது உணரத்தக்கது.

$$\chi^2_{\cdot 05\,(2)} = 5{\cdot}99$$

என அமைந்து. கணக்கிடப்பட்ட மதிப்பு எதிர்பார்க்கும் மதிப்பைவிட அதிகமாக இருத்தலின் கதிரலகிற்கும், விளைபலனுக்

கும் தொடர்பில்லை எனும் கருதுகோள் மறுக்கப்பட்டு, கதிரலகிற்கும் விளைபலனுக்கும் தொடர்புண்டு எனும் முடிவினைச் சொல்கிறோம்.

காட்டு :

180 பாத்திகளில் விதைகளை விதைப்பதற்காக எடுக்கப்பட்ட 180 ராண்டம் எண்களில் 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 அமைந்த நிகழ்வெண் பரவல் தரப்பட்டுள்ளது. இது ராண்டமாக அமைந்துள்ளது எனக் கொள்ள இடமளிக்கிறதா என்ற சோதனையை அமைப்போம்.

எண்	நிகழ்வெண்	எண்	நிகழ்வெண்
0	18	5	22
1	19	6	21
2	17	7	20
3	15	8	15
4	18	9	15

எடுக்கப்பட்ட மாதிரி ராண்டமாக இருக்கும்போது 180 ராண்டம் எண்களில் 0 முதல் 9 வரை உள்ள எண்கள் ஒவ்வொன்றும் சமமாக அமைந்திருக்கும் என எண்ணுகிறோம். அதாவது ஒவ்வோர் எண்ணும் 18 தடவை அமைய வேண்டும். நமது சோதனை முடிவுகள் இதற்குச் சற்று மாறுபட்டுள்ளன. கொடுக்கப்பட்ட சோதனை முடிவுகளுக்கும் நாம் எதிர்பார்க்கும் முடிவுகளுக்கும் உள்ள வித்தியாசம் குறிப்பிடத்தக்கபடி உள்ளதா என்பதே நமது சோதனையின் நோக்கமாகும்.

கொடுக்கப்பட்டுள்ள முடிவுகளுக்கும் நாம் எதிர்பார்க்கும் முடிவுகளுக்கும் (ஒவ்வோர் எண்ணும் 18 தடவை அமைய வேண்டும்) குறிப்பிடத்தக்க வித்தியாசம் இல்லை என்ற கருதுகோளை அமைத்து கை-வர்க்க மதிப்பைக் கணக்கிடுகிறோம்.

$$\chi^2 = \frac{(18-18)^2}{18} + \frac{(19-18)^2}{18} + \frac{(17-18)^2}{18} + \frac{(15-18)^2}{18}$$

$$+ \frac{(18-18)^2}{18} + \frac{(18-18)^2}{18} + \frac{(21-18)^2}{18}$$

$$+ \frac{(20-18)^2}{18} + \frac{(15-22)^2}{18} + \frac{(15-18)^2}{18}$$

$$= \frac{1}{18}\left[0 + 1 + 1 + 9 + 0 + 16 + 9 + 4 + 9 + 9\right]$$

$$= \frac{1}{18}[58] = 3\ 22$$

$$\chi^2{\cdot 05(9)} = 16{\cdot}92$$

நமது கருதுகோள் உண்மையாக இருக்கும்போது மாதிரி எடுத்தலின் ஏற்ற இறக்கங்களினுல் எதிர்பார்க்கும் கை-வர்க்கப் பட்டியல் மதிப்பு ·05 சிறப்பெல்லேயில், கட்டின்மை எண்ணிக்கை ஒன்பதிற்கு 16·92 என உள்ளது. நாம் கணக்கிட்ட கை-வர்க்க மதிப்பு (3·22) இதனின்றும் மிகக் குறைவாக உள்ளது. நமது கருது கோள் ஏற்கப்பட்டு ராண்டம் மாதிரியில் அமைந்த 0, 1, 2, 3, 4, 5, 6, 7, 8, 9-க்கான பரவலில் ஒவ்வோர் எண்ணும் 18-லிருந்து சற்று மாறுபட்டுள்ளது குறிப்பிடத்தக்க வித்தியாசத்தைத் தராமல், எடுக்கப்பட்ட மாதிரி ஒரு ராண்டம் மாதிரி எனக் கொள்ள இடமளிக்கிறது என்ற முடிவைச் சொல்கிருேம்.

காட்டு :

100 நபர்கள் கொண்ட 10 மாதிரிகளின் ஆண்/பெண் பகுப்புக் கொடுக்கப்பட்டுள்ளது. ஆண்/பெண் பகுப்பு 50 : 50 என அமையும் கருதுகோளேச் சோதனே செய்வோம்.

ஆண்	40	52	49	50	53	48	42	45	41	51
பெண்	60	48	51	50	47	52	58	55	59	49

நமது கருதுகோள் உண்மையாக இருக்கும்போது ஒவ்வொரு மாதிரியிலும் 50 : 50 என்ற முறையில் ஆண்/பெண் பகுப்பு முறை அமையவேண்டுமென எதிர்பார்க்கிருேம். ஒவ்வொரு

மாதிரியிலும் 50/50 என அமையவேண்டியதே நாம் எதிர் பார்க்கும் நிகழ்வெண்ணாகும். இது பத்து மாதிரிகளுக்கும் பொருந்தும்.

$$\chi^2 = \Sigma\frac{(O_i - E_i)^2}{E_i} = \frac{(40 - 50)^2}{50} + \frac{(60 - 50)^2}{50}$$

$$+ \frac{(52 - 50)^2}{50} + \ldots\ldots + \frac{(49 - 50)^2}{50}$$

$$= \frac{2}{50}\left[10^2 + 2^2 + 1^2 + 0 + 3^2 + 2^2 + 8^2 + 5^2 + 9^2 + 1^2 \right]$$

$$= \frac{329}{25} = 13{\cdot}16$$

நமது கருதுகோள் உண்மையாக இருக்கும்போது கட்டின்மை எண்ணிக்கை ஒன்பதிற்கான 0·05 சிறப்பெல்லை மதிப்பு 16·92 என அமைந்துள்ளதால், 10 மாதிரிகளும் 50: 50 என்ற முறையில் ஆண்/பெண் பகுப்பைக் கொண்டுள்ளது என்ற கருது கோளை ஏற்றுக்கொள்கிறோம். கொடுக்கப்பட்ட முடிவுகளுக்கும் நாம் எதிர்பார்க்கும் முடிவுகளுக்கும் உள்ள வேறுபாடு குறிப் பிடும்படி அமையாமல் மாதிரி எடுத்தலின் ஏற்ற இறக்கங் களினால் அமைந்தது எனக் கொள்ள இடமளிக்கிறது என்ற முடிவைச் சொல்கிறோம்.

மாறுபாட்டுக் குறியெண் (Indices of Dispersion), ஈருறுப்பு மாறு பாட்டுக் குறியெண்

எடுக்கப்பட்ட மாதிரியின் அளவு கருதுகோள் பரவல்களைப் பொருத்த முடியாமல் மிகக் குறைந்த உறுப்புகளைக் கொண்ட விடத்து பாய்ஸான் மற்றும் ஈருறுப்புப் பரவல்களுக்கான சில சோதனைகளை அமைப்போம்.

கருதுகோள் மாறுபாட்டளவையுடன் மாதிரியின் மாறு பாட்டளவை ஒத்து அமைகிறதா என்ற கருதுகோள் இம் மாதிரிச் சோதனைகளில் சோதனை செய்யப்படுகிறது. இணைப்புப் பட்டி யலுக்கான சோதனை முறையில் சிறு மாறுதல் செய்து இம் மாதிரி அளவுகள் சோதனை செய்யப்படுகின்றன.

x_1, x_2 .... x_k என்பன K மாதிரிகளில் n நிகழ்ச்சிகளில் கிடைத்த வெற்றிகளைக் குறிப்பதாகக் கொள்வோம். இம் மாதிரிகள் யாவும் ஈருறுப்புப் பரவலைச் சேர்ந்ததாகக் கொள்ளும் போது $(n - x_1)$, $(n - x_2)$ .... $(n - x^k)$ என்பன அவற்றுக்

கான தோல்வியைக் குறிக்கும். வெற்றிகளும் தோல்விகளும் $2 \times K$ எனும் இணைப்புப் பட்டியலாக அமைகிறது.

$$x_1,\ x_2,\ x_3 \ldots\ldots\ldots\ldots x_k.$$
$$n - x_1,\ n - x_2,\ n - x_3 \ldots\ldots\ldots n - x^k.$$

இதனை ஓர் இணைப்புப்பட்டியலுக்குச் சமமாக்கும்போது முதல் வரிசைக்கான நாம் எதிர்பார்க்கும் நிகழ்வெண்கள்.

$$e_i = \frac{\sum_1^k x_i}{nk} \times n = \bar{x}$$

$\bar{x}$ முதல் வரிசையின் எதிர்பார்க்கும் நிகழ்வெண்ணாகும்போது' துணைமொத்தங்கள் சமமாக இருக்கவேண்டும் எனும் கட்டளை இரண்டாவது வரிசைக்கான எதிர்பார்க்கும் எண்களை $(n - \bar{x}$ என அமைக்கிறது. இம் மதிப்புகளுக்கான கை-வர்க்க அளவு

$$\chi^2 = \sum_1^k \frac{(x_i - x)^2}{\bar{x}} + \sum_1^k \frac{(x_i^1 - \bar{x})^2}{n - \bar{x}}$$

$$= \left(\frac{1}{x} + \frac{1}{n - \bar{x}}\right) \Sigma\,(x_i - \bar{x})^2$$

$$= \frac{\sum^k (x_i - \bar{x})^2}{x\,(1 - \bar{x}/n)}$$

சாதாரணமாக அமைக்கக் கூடிய இணைப்புப் பட்டியலிலிருந்து இப் பட்டியல் சிறிது மாறுபட்டிருப்பது குறிப்பிடத்தக்கது. சாதாரணமாக அமைக்கும் இணைப்புப் பட்டியலில் அளவுகள் பட்டியலின் ஏதாவதோர் அறையில் அமையும் தன்மையையும், இப் பட்டியலில் முதல் n அளவுகள் முதல், இரண்டாவது வரிசையில் அமைந்துள்ள முதல் அறைகளில் ஏதாவதொன்றில் அமையும் நிலையில் உள்ள வேறுபாடு உணரத்தக்கது. இப் பரவல் $(k - 1)$ கட்டின்மை எண்ணிக்கையைக் கொண்டதாய் அமைகிறது.

கை-வர்க்க மதிப்பிற்கான வாய்பாட்டின் பகுதி மாதிரி மாறுபாட்டளவையின் k பங்கினைக் குறிக்கும். ஈருறுப்புப் பரவலின் மாறுபாட்டளவை $npq = np(1 - p) = \mu(1 - \mu/n)$-க்குச் சமமாகிறது. μ ஐ மாதிரியின் கூட்டுச் சராசரியாகக் கொள்ளும்போது

கை-வர்க்க மதிப்பீற்கான வாய்பாட்டின் தொகுதியும் மாறுபாட்டளவையின் மற்றொரு மதிப்பீடாக அமைகிறது. கை-வர்க்கம், இரண்டு மாறுபாட்டளவைகளின் விகிதத்தின் k-பங்கினைக் குறிக்கிறது. மாறுபட்ட முழுமைத் தொகுதியிலிருந்து மாதிரிகள் எடுக்கப்படும்போது, கை-வர்க்க அளவையின் பகுதி, தொகுதியைக் காட்டிலும் அதிக மதிப்புடன் அமைந்து குறிப்பிடும்படியான வித்தியாசத்தைச் சொல்லக் கூடியதாக அமையும். எனவே, இதன் முடிவும், கொடுக்கப்பட்ட அளவுகள் ஈருறுப்பு முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டதா என்பதற்குச் சமமாகிறது.

காட்டு:

90 செடிகள் கொண்ட 12 பாத்திகளில் பாதிக்கப்பட்ட செடிகளின் எண்ணிக்கை 19, 6, 9, 18, 15, 13, 14, 15, 16, 20, 22, 14 எனக் கொடுக்கப்பட்டுள்ளது. 12 பாத்திகளிலும் செடிகள் பழுதாவதற்கான நிகழ்திறன் சமமானதா என்பதற்கான சோதனையைப் பார்ப்போம். $p_1, p_2 \ldots\ldots p_{12}$ என்பன ஒன்று, இரண்டு... ஆகிய பாத்திகளில் பழுதாகும் செடிகளுக்கான நிகழ்திறனைக் குறித்துக்கொண்டால் நமது கருதுகோள் $p_1 = p_2 = p_3 = \ldots\ldots = p_{12}$ என அமைகிறது. 'இவற்றில் குறிப்பிடும்படியான வித்தியாசம் இல்லை' என்ற கருதுகோளை அமைத்துச் சோதனைச் செய்கிறோம்.

பாத்தி எண்	பழுதான செடிகள்	பாத்தி எண்	பழுதான செடிகள்
1	19	7	14
2	6	8	15
3	9	9	16
4	18	10	20
5	15	11	22
6	13	12	24

$$\bar{x} = \frac{181}{12} = 15 \cdot 1$$

$$\Sigma (xi - \bar{x})^2 = 228$$

$$\chi^2 = \frac{\Sigma (ni - \bar{x})^2}{\bar{x}(1 - \bar{x}/n)} = 18$$

கட்டின்மை எண்ணிக்கை பதினொன்றிற்கான 0·05 சிறப் பெல்லேயில் நாம் எதிர்பார்க்கும் கை-வர்க்க அளவு 19·7 என அமைந்துள்ளது. நமது கருதுகோள் ஏற்கப்பட்டுச் செடிகள் பழுதாகும் தன்மைக்கான நிகழ்திறன் எல்லாப் பகுதிகளுக்கும் சமமாய் உள்ளது என்ற கருத்தைச் சொல்கிறோம்.

பாய்ஸான் மாறுபாட்டுக் குறியெண்

p-ன் மதிப்பு மிகவும் குறைவாய் இருக்கும்போது, n அதிகப் படுத்தப்பட்டால் $\bar{x}/n$-மாதிரியின் p மதிப்பீடு மிகவும் குறைந்து $(1-\bar{x}/n)$ ஐ ஒன்றிற்குச் சமமாக்கிவிடும். ஈருறுப்பு மாறுபாட் டளவைக் குறியீட்டெண்ணில், இந்த அடிப்படையில் அமையும் எண்ணே பாய்ஸான் மாறுபாட்டளவைக் குறியீட்டெண் எனப் படும்.

$$\sum_{1}^{k}\frac{(x_i-\bar{x})^2}{\bar{x}}$$

பாய்ஸான் குறியீட்டெண், எதிர்பார்க்கும் நிகழ்வெண் சமமாக அமையும் கை-வர்க்க மதிப்பினுக்குச் சமமாகத் தோன்றினாலும் இதில் உள்ள சிறு மாற்றத்தை உணர்வதவசியம். சாதாரணமாக அமைக்கும் சோதனையில் நிகழ்வெண் கூட்டுத்தொகை 'n' மொத்த நிகழ்ச்சிகளின் கூட்டுத்தொகைக்குச் சமமாகும். ஆனால், பாய் ஸான் குறியீட்டெண்ணில், k அளவுகளும் கொடுக்கும் மதிப்பு ஒரே எண்ணாக அமைகிறது.

காட்டு ;

1000 விதைகள் கொண்ட 11 பொட்டலங்களில் பழுதான விதைகள் 15, 13, 8, 6, 11, 9, 14, 10, 16, 9, 12 என அமைந் துள்ளன. விதைகள் பழுதாவதற்கான நிகழ்திறன் 11 பொட் டலங்களிலும் சமமானதா என்பதற்கான சோதனையைக் காணப் பாய்ஸான் குறியீட்டெண்ணைக் கணக்கிடுவோம்.

$$\bar{x} = \frac{123}{11} = 11{\cdot}2$$

$$\Sigma(x-\bar{x})^2 = 97{\cdot}6$$

$$\chi^2 = \frac{\Sigma(x_i-\bar{x})^2}{\bar{x}} = \frac{97{\cdot}6}{11{\cdot}2} = 8{\cdot}7$$

$$\chi^2_{\cdot 05(10)} = 18{\cdot}3$$

விதைப் பொட்டல எண்	பழுதான விதைகள்	விதைப் பொட்டல எண்	பழுதான விதைகள்
1	15	7	14
2	13	8	10
3	8	9	16
4	6	10	9
5	11	11	12
6	9		

கொடுக்கப்பட்ட அளவுகள் பாய்ஸான் முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டவை எனக் கொள்வதில் தவறில்லை என்ற முடிவைச் சொல்கிறோம். விதைகள் பழுதாவதற்கான நிகழ்திறன் எல்லாப் பொட்டலங்களிலும் சமமாக அமைந்து, இது ஒரு பாய்ஸான் முழுமைத் தொகுதியைச் சேர்ந்தது எனக் கொள்கிறோம்.

பயிற்சி :

1. பட்டாணிபற்றிய ஆராய்ச்சி ஒன்றில் மெண்டல் அடியிற்காணும் முடிவுகளைக் குறித்துள்ளார். 315 விதைகள் வட்டமாகவும் மஞ்சள் நிறத்தைப் பெற்றதாயும், 101 விதைகள் திரைவுகள் நிறைந்ததும் (wrinkle) மஞ்சள் நிறத்தைப் பெற்றதாயும், 108 விதைகள் வட்டமாகவும் பச்சை நிறத்துடனும், 32 விதைகள் திரைவுகள் நிறைந்ததும் பச்சையாகவும் அமைந்துள்ளன. 9 : 3 : 3 : 1 என்ற முறைப்படி இந்த முடிவுகள் அமைய வேண்டுமெனக் கருதப்படுகிறது. கை-வர்க்க உதவிகொண்டு கருதுகோளுக்கும், சோதனை முடிவிற்கும் உள்ள ஒற்றுமையை விவரி.

2. மரபுவழிப் பண்பியல் உருப்படிவப்படி (genetic model) மூன்று பகுதிகளின் எண்ணிக்கை $p^2 : 2pq : q^2$ என அமைய வேண்டுமென எதிர்பார்க்கப்படுகிறது. இவை $p + q : 1$ என்ற கட்டளையை ஒத்து அமைகிறது. 9, 51, 45 என்ற சோதனை முடிவு, இம் மூன்று பகுதிகளும் $p^2 : 2pq : q^2$ எனும் மரபுவழிப் பண்பியல் உருவத்தை ஒத்து அமைந்துள்ளன எனக் கொள்ள முடியுமா என்பதைச் சோதனை செய்.

3. மரபுவழிப் பண்பியல் உருப்படிவ முறைப்படி ரத்தப் பகுப்புமுறை $O = q^2$, $A = p^2 + 2pq$, $B = r^2 + 2qr$, $AB = 2pr$ என அமைய வேண்டுமென எதிர்பார்க்கப்படுகிறது. இவை $p + q + r = 1$ எனும் கட்டளையை ஒத்து அமைகிறது. $O = 176$, $A = 182$, $B = 60$, $AB = 17$ என அமையும் சோதனை முடிவுகள் கொடுக்கப்பட்ட உருப் படிவத்தை ஒத்து அமைந்துள்ளதா என்பதற்கான சோதனையைச் செய்.

4. கொடுக்கப்பட்ட நிகழ்வெண் பரவலுக்கான மாறி பாய்ஸான் பரவலை ஒத்து அமைகிறது என்ற கருதுகோட்படி கணக்கிடப்பட்ட நிகழ்வெண்கள் தரப்பட்டுள்ளன. 98 விதைப் பொட்டலங்களில் பழுதான விதைகளின் எண்ணிக்கைக்கான சோதனை முடிவுகளை இச் சோதனை தருகிறது.

பழுதான விதை	0	1	2	3	4	5	6	7	8	9	10	≥ 10
சோதனை முடிவு	3	17	26	16	18	9	3	5	0	1	0	0
எதிர்பார்க்கும் முடிவு	4·78	14·44	21·81	21·96	1·58	10·02	5·04	2·18	0·82	0·27	·08	·01

'இது ஒரு பாய்ஸான் பரவலை ஒத்து அமைகிறது' என்ற கருதுகோளைச் சோதனைக் குள்ளாக்கு.

5. 500 அளவுகளுக்கான இயல்நிலைப் பரவலின் சோதனை முடிவுகள் மற்றும் எதிர்பார்க்கப்படும் முடிவுகள் தரப்பட்டுள்ளன. எதிர்பார்க்கப்படும் முடிவுகளைக் கணக்கிடக் கூட்டுச் சராசரி (x) திட்டவிலக்கம் (σ) சோதனை முடிவுகளிலிருந்து கணக்கிடப்பட்டுள்ள நிலையில், கொடுக்கப்பட்ட சோதனை முடிவிற்கான மாறி ஓர் இயல்நிலைப் பரவலை ஒத்து அமையும் என்பதற்கான சோதனையைச் செய்.

சோதனை முடிவு	5	12	43	61	105	103	89	54	19	7	2
எதிர் பார்க்கப் படும் முடிவு	5	14	36	71	102	101	85	50	21	7	2

6. ஒரு குறிப்பிட்ட மருந்து தடிமல் (cold) தடுப்பில் சக்தி வாய்ந்தது எனக் கருதப்படுகிறது. தடிமல் பாதிப்புடைய 164

நபர்களில் 82 பேருக்குக் குறிப்பிட்ட மருந்தையும், 82 பேருக்குச் சர்க்கரைக் கலவையையும் கொடுத்துக் கீழ்க்கண்ட முடிவுகள் பெறப்படுகின்றன.

	குணமானவர்கள்	பாதிக்கப் பட்டவர்கள்	எந்தக் குறிப்பும் தெரியாதவர்கள்	மொத்தம்
மருந்து சாப்பிட்டவர்கள்	52	10	20	82
சர்க்கரைக் கலவை சாப்பிட்டவர்கள்	44	12	26	82
	96	22	46	164

'சர்க்கரைக் கலவையைவிட மருந்து சிறந்ததில்லை' எனும் கருதுகோளை இச் சோதனை முடிவுகளைக் கொண்டு முடிவு செய்.

7. கார்ல் பியர்ஸன் (Karl Pearson) வாழ்க்கை நிகழ்ச்சிக் கோவையில் (memoir) அடியிற்காணும் குறிப்புக் கொடுக்கப்பட்டுள்ளது.

		பையன்களின் கண் நிறம்		
		ஒளி உடையவை	ஒளி குறைந்தவை	மொத்தம்
தந்தை கண் நிறம்	ஒளி உடையவை	230	148	378
	ஒளி குறைந்தவை	151	471	622
	மொத்தம்	381	619	1000

பையன்களின் கண் நிறம் தந்தைமார்களின் கண்ணிறத்தை ஒத்து அமைய இடமிருக்கிறதா எனும் கருதுகோளைச் சோதனை செய்.

8. ஒருவிதமான தொத்துநோய்ப்(epidemic) பரவலில் 927 குழந்தைகள் பாதிப்புக்குள்ளானார்கள். இவர்களில் 408 பேர் மருத்துவப் பண்டுவம் (treatment) பெறுதவர்களாகவும், 104 பேர் நோய் தொற்றியபின் சிகிச்சை செய்யப்பட்டவர்களாகவும் குறிக்கப்பட்டுள்ளனர். மீதமுள்ள மருத்துவப் பண்டுவம் பெற்றவர்களில் 166 பேர் தொற்றுநோய்க்குள்ளாகி உள்ளனர். 'மருத்துவப் பண்டுவம் இவ் வியாதியைத் தடுப்பதில் சக்தி வாய்ந்ததில்லை' எனும் கருதுகோளைச் சோதனை செய்.

9. ஆடுகளின் நோய் ஒன்றிற்கான தடுப்பூசி போட்ட முயற்சியில் அடியிற்காணும் முடிவுகள் பெறப்படுகின்றன. தடுப்பூசி போடுவதற்கும் நோய் வராமல் தடுப்பதற்கும் தொடர்புண்டா என்பதைச் சோதனை செய்.

	இறந்தது	குணமானது	மொத்தம்
தடுப்பூசி போட்டவை	2	10	12
தடுப்பூசி போடாதவை		6	12
	8	16	24

யேட்சின் திருத்தத்துடன் (Yates correction) கை-வர்க்க மதிப்பைக் கணக்கிடு.

10. ஒன்பது நாணயங்களை 512 தடவை வீசியதில் கிடைத்த முடிவுகள் தரப்பட்டுள்ளன.

தலைகளின் எண்ணிக்கை	0	1	2	3	4	5	6	7	8	9
நிகழ் வெண்	1	2	1	50	154	110	100	92	1	1

பழுதற்ற நாணயங்களாக அமையப்பெற்றது என்ற அடிப்படையில் நாம் எதிர்பார்க்கும் தலைகளுக்கான ஈருறுப்புப் பரவல் $512\left(\frac{1}{2}+\frac{1}{2}\right)^9$ என அமைகிறது. இக் கருதுகோட்படி நாம் எதிர்பார்க்கும் முடிவுகள்,

தலைகள்:	0	1	2	3	4	5	6	7	8	9
நிகழ்வெண்:	1	9	36	84	126	126	84	36	9	1

எனக் கணக்கிடப்பட்டுள்ளது.

'நாணயங்கள் பழுதற்றவை (unbiased) எனும் கருதுகோட்படி முடிவுகள் அமைந்துள்ளன' என்பதைச் சோதனை செய்.

11. குற்றவாளிகள் மனநிலைக்கும் எடைக்கும் தொடர்புண்டா எனும் கருதுகோளைச் சோதனை செய்ய எடுக்கப்பட்ட மாதிரியில் அடியிற்காணும் முடிவுகள் தரப்பட்டுள்ளன.

	எடை				
மனநிலை	90–120	120–130	130–140	140–150	150–
இயல்பானது (Normal)	21	51	94	106	125
குறையுடையது (Weak)	15	18	34	15	15

சோதனை முடிவுகள் குற்றவாளிகள் மனநிலைக்கும் எடைக்கும் தொடர்புண்டு என்ற நிலையைக் கூறுகிறதா ?

12. 200 ராண்டம் எண்கள் எடுத்ததில் அடியிற்காணும் முடிவுகள் பெறப்படுகின்றன.

எண்	0	1	2	3	4	5	6	7	8	9
நிகழ்வெண்	18	19	23	21	16	25	22	20	21	15

எல்லா எண்களும் சமமாகப் பரவியுள்ள பட்டியலிலிருந்து இம் மாதிரி எடுக்கப்பட்டதா என்பதற்கான சோதனையைச் செய்.

13. ஒவ்வொரு மாதிரியிலும் 500 விதைகள்கொண்ட பதினோர் இயைபிலா மாதிரிகளில் ஒரு குறிப்பிட்ட விதை

25 25 27 28 31
25 27 28 30 33 33

என அமைந்துள்ளது. ஈருறுப்பு மாறுபாட்டுக் குறியெண்ணைக் கணக்கிட்டு இவை ஓர் ஈருறுப்புப் பரவலை ஒத்து அமைந்துள்ளதா என்பதற்கான சோதனையைச் செய்.

14. 4 குழந்தைகள் உள்ள 800 குடும்பங்களில் ஆண்/பெண் பிரிவினுக்கான சோதனை முடிவுகள் தரப்பட்டுள்ளன.

ஆண் குழந்தைகள்	பெண் குழந்தைகள்	குடும்பங்கள்
0	4	32
1	3	178
2	2	290
3	1	236
4	0	64

ஆண் குழந்தை பிறப்பதற்கும் பெண் குழந்தை பிறப்பதற்கும் உள்ள வாய்ப்புச் சமமானது என்ற கருதுகோளைச் சோதனை செய்.

15. அணு உயிர்க்குழு (bacterial colony) ஒன்றின் வளர்ச்சிக்கான முடிவுகள் தரப்பட்டுள்ளன. 15 தட்டுகளில் குறிப்பிட்ட கலவையில் வளர்க்கப்பட்ட முடிவுகளை இவை குறிக்கின்றன. சோதனை முடிவுகள் உயிரணுக்கள் கலவையில் ராண்டமாகப் பரவி இருந்தது என்ற கருதுகோளை உறுதிசெய்கிறதா என்பதற்கான சோதனையைச் செய். உயிரணுக்களின் எண்ணிக்கை :

193	153	171	140	133
168	183	156	151	164
161	152	159	152	157

14. மாறுபாட்டளவை ஆய்வு

(Analysis of Variance)

தொகுக்கப்பட்ட செய்திகளைப் பல பகுதிகளாகப் பிரிக்கும் போது அவற்றுக்கான மாறுபாட்டின் தன்மையைப் படிப்பதே மாறுபாட்டின் ஆய்வு எனப்படும். திட்டவிலக்கத்தின் வர்க்கமே மாறுபாட்டளவை என்பதை அறிவோம். மாறியின் தனிப்பட்ட மதிப்புகளிலிருந்து, அவற்றின் கூட்டுச் சராசரியைக் கழித்து வரும் விலக்கங்களுக்கான $(x_i - \bar{x})$ வர்க்க மொத்தத்தின் சராசரி அளவு மாறுபாட்டளவையைக் குறிக்கும். இது மாறியின் மொத்த மாறுபாட்டளவை (total variance) எனப்படும். இதனை மதிப்பிடுவதே மாறுபாட்டளவையைக் கணக்கிடுவதன் நோக்கமாக அமைந்துள்ளது. மொத்த அளவுகள் சில பகுதிகளாகப் பிரிக்கப் படும்போது, மொத்த மாறுபாட்டளவையில் இப் பகுதிகளின் பங்கு எவ்வளவு என்பதையும், இப் பகுதிகளுக்கான மாறுபாட் டளவைகள் குறிப்பிடும்படியான வித்தியாசத்துடன் அமைந் துள்ளதா என்பதைக் கணக்கிடுவதும் மாறுபாட்டளவை ஆய்வின் மற்றொரு நோக்கமாகும். முழுமைத் தொகுதியின் பல பகுதி களுக்கான கூட்டுச் சராசரிகள் குறிப்பிடும்படியான வித்தியாசத் தைக் கொண்டுள்ளனவா என்பதைச் சோதனை செய்வதையே இது குறிக்கும். t-பரவலின் உதவியால் இரண்டு கூட்டுச் சராசரி களுக்கான வித்தியாசம் குறிப்பிடும்படி உள்ளதா என்பதை ஆராய்ந்தோம். இங்கு இரண்டு அல்லது அதற்கு மேற்பட்ட கூட்டுச் சராசரிகளின் வித்தியாசத்திற்கான சோதனை அமையும் சிறப்பைக் காண்கிறோம்.

நான்கு இனப் பசுக்களடங்கிய பண்ணையிலிருந்து ராண்டம் மாதிரி ஒன்றெடுப்பதாகக் கொள்வோம். 36 பசுக்கள் அடங்கிய இம் மாதிரியில் 4 இனப் பசுக்களும் சமமாக அமைந்துள்ள தாகவும் கொள்வோம். இப் பசுக்கள் யாவற்றிற்கும் ஒரு குறிப்

பிட்ட காலத்திற்கான பால் கொடுக்கும் அளவு குறிக்கப் பட்டுள்ளது. 36 பசுக்களின் தினசரிப் பால் அளவிற்கான பரவல் ஒரு மாறுபாட்டளவையைக் கொடுக்கும். 4 இனப் பசுக்களும் தனித்தனியாகப் பிரிக்கப்பட்டு 4 பகுதிகளாக அமைக்கப்படும் போது, ஒவ்வோர் இனத்திற்குமான மாறுபாட்டளவைகள் கிடைக்கப்பெறுவோம். இவற்றுக்கான மாறுபாட்டளவைகள் குறிப்பிடும்படியான வித்தியாசத்துடன் அமைந்துள்ளதா என்ப தையும், 4 இனப் பசுக்களின் கூட்டுச் சராசரிகளின் அளவு 36 பசுக்களுக்குமான மொத்தக் கூட்டுச் சராசரியினின்றும் மாறு பட்டுள்ளதா என்பதையும் மாறுபாட்டளவை ஆய்வில் சோதனை செய்கிறோம். மாறுபாட்டளவை ஆய்வின்போது இரு சோதனை கள் செய்யப்படுவதை உணர்வதவசியம். மொத்த மாறுபாட்டிற் கான பகுதிகளின் மாறுபாட்டளவைகளைக் கொடுப்பதையும், சிறப்புக்காணும் சோதனை அமைந்துள்ளதையும் காணலாம்.

F-பரவல் (F- distribution)

S_1^2, S_2^2 என்பன இரண்டு மாதிரிகளுக்கான மாறுபாட் டளவைகளாக இருக்கும்போது, முழுமைத் தொகுதியின் மாறு பட்டளவையின் பிறழ்ச்சியற்ற மதிப்பீடாக

$$U_1^2 = \frac{n_1 S_1^2}{n_1 - 1}$$

$$U_2^2 = \frac{n_2 S_2^2}{n_2 - 1}$$

என்பன அமைவதை அறிவோம். $n_1 - 1 = \delta_1$, $n_2 - 1 = \delta_2$ என்பன மதிப்பீடு செய்வதற்கான கட்டின்மை எண்ணிக்கை களாகும். U_1^2 மற்றும் U_2^2 இரண்டிற்குமான வித்தியாசம் குறிப்பிடும்படி அமைந்துள்ளதா என்பதற்கான சோதனையே நமது கருதுகோளாகும். இவற்றுக்கான வித்தியாசம் குறிப் பிடும்படி அமையாது, இவை இரண்டும் ஒரே மாறுபாட்டளவை (same variance) கொண்ட முழுமைத் தொகுதியிலிருந்து எடுக் கப்பட்டனவா என்பது சோதனையின் தன்மையாகும்.

$$F = \frac{U_1^2}{U_2^2} = \frac{n_1 S_1^2/n_1 - 1}{n_2 S_2^2/n_2 - 1} \quad (U_1^2 > U_2^2)$$

இதன் அளவை δ_1, δ_2 கட்டின்மை எண்ணிக்கைகளுடன் F என்ற மாதிரியின் பரவலாக (sampling distribution of F) அமைகிறது. $U_1^2 = U_2^2$ என அமையும்போது $F = 1$ என அமைகிறது. F-பரவலின் பகுதி (U_1^2) எப்போதும் F-பரவலின் தொகுதியைக்

$(U_2{}^2)$ காட்டிலும் அதிகமாக இருக்கும்படி அமைப்பதால் நம்மால் கணக்கிடப்படும் மதிப்பு ஒன்றிலிருந்து எவ்வளவு விலகி இருக்கிறது என்பதைக் குறிக்கும்.

நமது கருதுகோள் உண்மையாக இருக்கும்போது, மாதிரியின் ஏற்ற இறக்கங்களால் ஏற்படக்கூடிய F-பரவலின் ·05, ·01 சிறப்பெல்லைக்கான நிகழ்திற மதிப்பு F-பட்டியலில் கொடுக்கப் பட்டுள்ளது.

F-பட்டியல்

கட்டின்மை எண்ணிக்கை δ_2 \ கட்டின்மை எண்ணிக்கை δ_1		1	2	3	....	200	500	∞
1	5%							
	1%							
2	5%							
	1%							
3	5%							
	1%							
· · · · · ·								
∞	5%							
	1%							

பட்டியலில் ஒரு மதிப்பைப் பார்க்கும்போது, F- அளவைக்கான வாய்பாட்டின் பகுதியில் அமைந்த மாறுபாட்டளவைக்கான கட்டின்மை எண்ணிக்கையைப் படுகை வரிசையிலும் தொகுதியில் அமைந்த மாறுபாட்டளவையின் கட்டின்மை எண்ணிக்கையை நிமிர்நிலை வரிசையிலும் குறித்து, இவை இரண்டும் சந்திக்கும் அறையில் அமைந்த 0·05, ·01 சிறப்பெல்லைக்கான மதிப்பைக் குறித்துக்கொள்ள வேண்டும். கணக்கிடப்பட்ட மதிப்பு நாம் எதிர்பார்க்கும் மதிப்பைவிடக் குறைவாக இருந்தால் கருதுகோள் ஏற்கப்பட்டும், அதிகமாக இருந்தால் கருதுகோள் மறுக்கப்பட்டும் முடிவு செய்யப்படும்.

காட்டு :

10, 8 எலிகள்கொண்ட இரு மாதிரிகளின் எடை உயர்வுக்கான சோதனையில்

$$n_1 = 10 \qquad S_1^2 = 10\ 24$$
$$n_2 = 8 \qquad S_2^2 = 10 \cdot 25$$

எனக் கொடுக்கப்பட்டுள்ளது. இவ்விரண்டு மாதிரிகளுக்குமான மாறுபாட்டளவையில் குறிப்பிடும்படியான வித்தியாசமின்றி இரண்டும் ஒரே மாறுபாட்டளவை உடைய முழுமைத் தொகுதியைச் சேர்ந்ததெனக் கொள்ளலாமா என்ற சோதனையைச் செய்வோம்.

கொடுக்கப்பட்ட மாதிரிகளின் மாறுபாட்டளவைகளிலிருந்து, முழுமைத் தொகுதிக்கான பிறழ்ச்சியற்ற மதிப்பீடுகளை அமைத்து F-ன் மதிப்பைக் கணக்கிடுகிறோம்.

$$U_1^2 = \frac{n_1 S_1^2}{n_1 - 1} = 11\ 38$$

$$U_2^2 = \frac{n_2 S_2^2}{n_2 - 1} = 11 \cdot 43$$

$$F = \frac{n_2 S_2^2 / n_2 - 1}{n_1 S_1^2 / n_1 - 1} = \frac{11 \cdot 43}{11 \cdot 38}$$

$$= \frac{U_2^2}{U_1^2} = 1 \cdot 005 \quad [\, U_2^2 > U_1^2]$$

பட்டியலில் கட்டின்மை எண்ணிக்கை 7,9-க்கான ·05 சிறப்பெல்லையின் மதிப்பு 3·29 என அமைந்துள்ளது. கணக்கிடப்பட்ட மதிப்பு நாம் எதிர்பார்க்கும் மதிப்பைவிடக் குறைவாக

இருப்பதால் இரண்டு மாறுபாட்டளவைகளுக்குமான வித்தியாசம் குறிப்பிடும்படி அமையாமல், இரண்டு மாதிரிகளும் ஒரே முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்டது எனக் கொள்ள இடமிருக்கிறது என்ற முடிவைச் சொல்கிறோம்.

செய்முறைத் திட்ட அடிப்படைக் கருத்துகள்

F-சோதனையின் அடிப்படைக் கருத்துகள் உயிர்ப்புள்ளியியலுக்கு மிக உபயோகமான 'செய்முறைத் திட்ட' (design of experiments) முடிவுகளை ஆராய்வதற்குப் பெரிதும் பயன்படுகின்றன. சில முக்கியமான செய்முறைத் திட்டங்களின் ஆய்வைக் காண்பதற்குமுன் செய்முறைத் திட்டத்திற்கான சில அடிப்படைக் கோட்பாடுகளைக் காண்போம்.

புள்ளியியல் முறைகளைச் சரியாகப் புரியாத சிலர், தாங்கள் எதிர்பார்க்கும் முடிவுகளைச் சோதனைமூலம் பெறமுடியாத விடத்து, புள்ளியியலார்தம் உதவியை நாடுவதைக் காண்கிறோம். சில சோதனைகளிலிருந்து எடுக்கப்பட்ட அளவுகள் புள்ளியியல் சோதனைகளின் தன்மையை மனத்தில் கொண்டு சேகரிக்கப்படாத நிலையிலும், சில அளவுகள் புள்ளியியல் முறைப்படி எந்த முடிவும் எடுக்கப்படாத நிலையிலும் அமைந்துள்ளன. செய்முறைத் திட்டங்களின் கோட்பாடுகளைச் சரிவரக்கொண்ட சோதனைகளின் முடிவுகளைக் காண்பது மிக அரிதாக உள்ளது. இக் கோட்பாடுகள் யாவும் செய்முறைத் திட்டங்களை அமைப்பதற்கு முன்பே மனத்தில் கொள்ள வேண்டியவையாகும். இவற்றை மறந்தும், மறுக்கப்பட்டும் செய்த செய்முறைத் திட்டங்களிலிருந்து கிடைத்த செய்திகள், புள்ளியியல் முறைப்படி ஆய்வு செய்ய முடியாதனவாகவோ நம்ப முடியாத முடிவுகளைக் கொடுப்பனவாகவோ இருக்கும்.

நமது செய்முறைத் திட்டத்தின் நோக்கத்திற்கான ஒரு மாறியின் அளவுகளிலிருந்து செம்மையான முடிவுகளை எடுக்க இம் மாறியின் அளவுகளைப் பாதிக்கும் பல்வேறுபட்ட தன்மைகளைக் கட்டுப்பாடு (control) செய்ய வேண்டியவர்களாக உள்ளோம். ஆராய்ச்சி நிலைய அமைப்பால் சில தன்மைகள் கட்டுப்படுவனவாயும், புள்ளியியல் செய்முறைத் திட்டத்தால் (design of experiments) சில தன்மைகள் கட்டுப்படுத்தப்படுவனவாயும் அமைகின்றன. நிலத்தை இரு சமபாகங்களாகப் பிரித்து இரண்டு இன விதைகளை விதைப்பதாகக் கொள்வோம். ஒவ்வொரு பிரிவிலும் ஓர் இன விதைகளை விதைத்து முடிவுகள் குறிக்கப்படுகின்றன. நிலத்தின் தன்மையில் குறிப்பிடும்படியான

வேறுபாடு உள்ள இடத்தில் விளைச்சலில் உண்டான வித்தியாசம் விதைகளின் தன்மையைச் சார்ந்தது எனக் கொள்ள முடியாதவர்களாக உள்ளோம்.

இயைபிலாத் தன்மை (randomization), திரும்பச் செய்தல் (replicate) ஆகிய செய்முறைத் திட்டக் கோட்பாடுகளைச் செயல்படுத்துதல் மூலம், செய்முறைத் திட்ட முடிவுகளின் நம்பகத் தன்மையை அதிகப்படுத்த முடியும். நமக்குக் கொடுக்கப்பட்டுள்ள நிலத்தைச் சம அளவு உள்ள பல பாத்திகளாகப் பிரித்து, ஒரு பாதிப் பாத்திகளில் ஒரு இன விதைகளையும் மற்றப் பாதிப் பாத்திகளில் மற்ற இன விதைகளையும் விதைப்பதால் நிலத்தில் உள்ள மாறுபாட்டின் தன்மை இரு இன விதைகளுக்கும் சமமாக்கப்படுவதை உணர முடிகிறது. ஒவ்வொரு பாத்தியிலும் விதைக்கப்படும் விதைகள் இயைபில்லாமல் எடுக்கப்படும்போது விளைச்சலில் உள்ள வித்தியாசம் விதைகளின் தன்மைக்கானது என்ற முடிவைச் சொல்ல நமக்கு உறுதி கிடைக்கிறது. இயைபிலாத் தன்மைமட்டும் நம்பகமான முடிவுகளை எடுக்க உதவும் போதுமான நியதி எனக் கொள்ள முடியாது. இரண்டு விதைகளையும் விதைப்பதற்கான பகுதிகளைத் தேர்ந்தெடுக்க ஒரு நாணயத்தைச் சுண்டித் தலை விழும்போது முதல் பகுதியில் இரண்டாவது இன விதைகளையும், பூ விழும்போது முதல் பகுதியில் முதல் இன விதைகளையும் விதைப்பதாகக் கொள்வோம். விதைகளுக்கான பாகம் அல்லது ஒரு பகுதிக்கான விதைகள் இயைபில்லாமல் எடுக்கப்படுவதை உணர்கிறோம். இவ்விடத்து இரண்டு இன விதைகளும் சம விளைதிறன் (equally productive) உடைத்தாய் இருந்தாலும் நிலத்தில் உள்ள வேறுபாடு விளைச்சலில் ஏற்பட்ட மாறுபாட்டை விதைகளுக்கானது என உறுதியாகச் சொல்ல முடியாத நிலையில் உள்ளோம். எனவே, சோதனையில் இயைபிலாத் தன்மையைக் கொண்டு மட்டும் உறுதிப்பாடான முடிவிற்கு வித்திட முடியாதென்பது தெளிவாகிறது. விளைநிலத்தைச் சிறுசிறு பகுதிகளாகப் பிரித்து, பிரிக்கப்பட்ட பாத்திகளில் இரண்டு இன விதைகளையும் இயைபிலாத் தன்மையுடன் விதைக்கும்போது விளைநிலத்தின் வேறுபாடு இரண்டின விதைகளுக்கும் சமமாக்கப்படுவதை உணர்கிறோம். விளைநிலம் பிரிக்கப்படும் தன்மை, விதைகளை இயைபிலாமல் விதைக்கும்போது நிலத்தின் வேறுபாடு இரண்டின விதைகளுக்கும் முடிந்தவளவிற்குச் சமமாக இருக்கும்படி அமைக்க வேண்டும். இம் முறைப்படி பல சோதனைப் பாத்திகள் அமைத்து ஒவ்வோர்

இன விதைகளும், ஒரு பாகத்தில் விதைப்பதற்கு மாருகப் பல பாத்திகளில் விதைக்க ஏதுவாய் அமையும் தன்மையே திரும்பச் செய்தல் (replicate) எனப்படும். சோதனை முடிவுகளின் நம் பகத் தன்மை அதிகப்பட அமைவதற்கு, இயைபிலாத் தன்மை யுடன் (randomization) தேவையான அளவிற்கு திரும்பச் செய் தலும் (replicate) நலம்.

இயைபிலாத் தன்மையும், திரும்பச் செய்தலும் சிறந்த செய்முறைத் திட்டங்களை அமைப்பதற்கு மட்டும் உதவுகின்றன எனக் கொள்ள இடமில்லை. புள்ளியியல் யுத்திகளைக் கொண்டு எடுக்கப்படும் முடிவுகளை நம்பி உள்ள சில சோதனைகளுக்கு இவை தேவையானது எனவும் உணர வேண்டும். நம்மால் அமைக்கப்பட்ட பெரும்பாலான சோதனை அளவைகள் இயை பிலா மாதிரியைத் தழுவி உள்ளதை அறிவோம். அவற்றைப் பயன்படுத்தச் செய்முறை அளவுகள் யாவும் இயைபிலாத் தன்மைத்தாய் அமைவது அவசியமாகிறது. ஓர் அளவின் மாறு பாட்டைத் தெரியக் குறைந்தது இரண்டளவைகளாவது தேவைப்படுவதால், திரும்பச் செய்தலும் ஒரு தேவையே எனக் கொள்வதில் தவறில்லை. இரண்டின விதைகளின் விளைச்சலில் உள்ள வித்தியாசத்தைக் கணக்கிட t-பரவலை உபயோகப் படுத்த வேண்டியவர்களாக உள்ளோம். t-பரவல் இயைபிலா மாதிரியின் பரவலாகவும், இதன் அளவைக் கணக்கிட மாதிரி யின் மாறுபாட்டளவை தேவைப்படுவதாலும் செய்முறை முடிவுகள் இயைபிலாத் தன்மைத்தாயும், திரும்பச் செய்யப் பட்ட அளவுகளுடனும் அமைக்கப்படவேண்டிய தேவை ஏற்படுகிறது.

இயைபிலாத் தன்மை, திரும்பச் செய்தல் ஆகிய யுத்திகள் பிறழ்ச்சியற்ற முடிவுகளைக் கொடுக்கும் தன்மைக்கு உபயோகப் பட்டாலும், நிலத்தின் வளமையில் உள்ள வேறுபாடு (difference in soil fertility) விளைச்சலுக்கான வேறுபாட்டில் உண் டாக்கிய விளைவினைக் (effect) கட்டுப்படுத்துவதாகக் கொள்ள முடியாது. நிலத்தின் தன்மையில் வேறுபாடு அதிகமாகும்போது விளைச்சலுக்கான வேறுபாடும் அதிகமாகிறது. விளைச்சலுக் கான வித்தியாசங்களைக் கணக்கிட t-பரவலை உபயோகப் படுத்தும்போது நிலவள வேறுபாடு அதிகமாக இருந்தால், நிலவள வேறுபாடு இல்லாத இடத்தைக் காட்டிலும் அதிகமான மாதிரி அளவு அமையப்பெற வேண்டியதாய் முடிகிறது. t-அளவையின் தொகுதியில் மாறுபாட்டின் மதிப்பீடு அமை வதால் இந் நிலை ஏற்படுகிறது. ஒரே தன்மைத்தான பாத்திகளை

அமைப்பதால் இச் சோதனை முறையின் கூருணர்வுத் திறத்தை (sensitiveness) அதிகப்படுத்தலாம். இக் கட்டுப்பாட்டை ஏற்படுத்துவது எளிதன்று. பாத்திகளை முடிந்தவளவிற்கு ஒரே தன்மைத்தாய் அமைத்து, புள்ளியியல் செய்முறைத் திட்டத்தைக் கொண்டு (statistical design) மாறுபாட்டின் தன்மைக்கான விளைவுகளைக் கணிசமாகக் கட்டுப்படுத்த முடியும். இங்குக் காண உள்ள செய்முறைத் திட்டங்கள் யாவும் கூருணர்வுத் திறம் கொண்ட செய்முறைத் திட்டங்களாகும். சில முக்கியமான அடிப்படைத் திட்டங்களையும் அவற்றின் ஆய்வுமுறை பற்றியும் பார்ப்போம்.

செய்முறையின் கூருணர்வுத் திறத்தை அதிகப்படுத்த உதவும் முக்கியமான யுத்தி, நம் சிக்கலுக்குரிய மாறியின் மொத்த மாறுபாட்டளவையை (total variance), சோதனையின் தன்மைக்கேற்ப பலப் பகுதிகளாகப் பிரிக்கும்போது, ஆய்விற்குரிய மாறியின் தன்மையில் குறுக்கிடும் மாறிகளுக்கான மாறுபாட்டளவுகளின் மதிப்பைக் குறைத்து, திட்டத்தின் கூருணர்வுத் திறத்தை அதிகப்படுத்துகிறோம். மாறுபாட்டளவை ஆய்வு, செய்முறைத் திட்டத்தில் மிகப் பயனுடையதாய் அமைவதற்கு இத் தன்மையே காரணமாய் அமைந்துள்ளது.

செய்முறைத் திட்டத்தைச் செயலாக்குவதன் நோக்கமாக சில கருதுகோட் சோதனை செய்பவர்களாகவோ அல்லது முழுமைத் தொகுதியின் அளவைகளை மதிப்பீடு செய்பவர்களாகவோ செயல்படுவோம். இவ் வடிப்படை நோக்கத்திற்கான கூருணர்வுத் திறம் வாய்ந்த செய்முறைத் திட்டங்களை அமைக்க மாறுபாட்டாய்வு உதவுவதாகக் கொண்டாலும் இதன் தத்துவம் கருதுகோட் சோதனைகளினாலேயே விளக்கப்படும்.

இயல்நிலைப் பரவலை ஒத்து அமையும் முழுமைத் தொகுதியிலிருந்து n உறுப்புகள்கொண்ட மாதிரியை எடுப்பதாகக் கொள்வோம். இவ் வுறுப்புகளைச் சோதனைகளின் தன்மைக்கேற்ப வகைப்படுத்தமுடியும் என்பதைச் சொன்னோம். இவ் வளவுகள் விளைபலனாக உள்ளபோது விதைகளின் இனத்திற்கேற்பவோ உரங்களின் தன்மைக்கேற்பவோ பிரித்தமைக்கலாம். இவை பாலின் அளவைக் குறிக்குமிடத்து, பசுக்களின் இனத்திற்கேற்பவோ ஒரு மாநிலத்தின் மாவட்டங்களுக்கேற்பவோ பிரித்தமைக்கலாம். x_{ij} என்பது j மாவட்டத்தில் i பசுவினது பாலின் அளவைக் குறிப்பதாகக் கொள்வோம். x_{ij} என்பதன் கீழ்க் குறிகளான (suffix) 'i' என்பது பசுவின் எண்ணையும் 'j'

என்பது மாவட்டத்தின் பெயரையும் குறிப்பதாகக் கொள்கிறோம். n_i என்பது 'i' பிரிவில் அமையும் எண்களைக் குறிப்பதாகவும், $\bar{x}_i$ என்பது அப் பிரிவிற்கான கூட்டுச் சராசரியையும் $\bar{x}$, N என்பன முறையே மொத்தத்திற்கான கூட்டுச் சராசரியையும், மொத்த உறுப்புகளையும் குறிப்பதாகக் கொள்வோம்.

$$N\bar{x} = \sum_i \sum_j x_{ij}$$

$$N_i \bar{x}_i = \sum_j x_{ij}$$

$$\sum_{ij} (x_{ij} - \bar{x}) = 0$$

$$\sum_j (x_{ij} - \bar{x}_i) = 0$$

கூட்டுத் தொகைக்கான (N உறுப்புகளின்) விலக்கவர்க்கத்தை அமைக்க, i பிரிவில் அமைந்த உறுப்புகளுக்கான விலக்கவர்க்கங்களைக் கணக்கிடுகிறோம்,

$$\sum_j (x_{ij} - \bar{x})^2 = \sum_j (x_{ij} - \bar{x})^2 + n_i (x_i - \bar{x})^2$$

எல்லாப் பிரிவுகளுக்குமான இதன் மதிப்பைக் கூட்டும்போது நமக்குத் தேவையான வர்க்கங்களின் கூட்டுத்தொகை கிடைக்கிறது.

$$\sum_i \sum_j (x_{ij} - x)^2 = \sum_i \sum_j (x_{ij} - \bar{x}_i)^2 + \sum_i n_i (\bar{x}_i - \bar{x})^2$$

இஃது ஒரு கணிதச் சமன்பாடாகும். எனவே, முழுமைத் தொகுதி இயல்நிலைப் பரவலை ஒத்து அமையா இடத்தும் இவ் வாய்பாடு உபயோகப்படுத்துவது குறிப்பிடத்தக்கதாகும்.

ஒருகட்டளைவிதிப் பிரிவிற்கான மாறுபாட்டு ஆய்வுமுறை (Analysis of variance for one creterion of classification)

மாதிரி எடுக்கப்படும் முழுமைத் தொகுதியில், நம்மால் பிரிக்கப்படும் பண்பின் தன்மை ஒருபடித்தாய் (homogeneous) இருப்பதாகக் கொள்வோம். பிரிக்க உதவும் பண்பு, மாறியின் அளவில் எந்த விளைவையும் உண்டாக்குவதில்லை எனக் கொள்கிறோம். பல பிரிவுகளாக அமைந்த இப் பிரிவுகளின் பண்பு ஒரே தன்மைத்தாய் அமையும். குறிப்பாக இவ்வொரு பிரிவும் முழுமைத் தொகுதிக்கான கூட்டுச் சராசரியும், மாறுபாட்டளவையும் கொண்ட பிரிவுகளாக அமையும். மொத்த விலக்க வர்க்கத்

திற்காக அமைந்த வாய்பாட்டில் உள்ள மூன்று தனிப்பட்ட விலக்கவர்க்கங்களும்

$$\sum_i \sum_j (x_{ij} - \bar{x})^2 = \sum_i \sum_j (x_{ij} - \bar{x}_i)^2 + \sum n_i (x_i - \bar{x})^2$$

மாறுபாட்டளவைக்கான மூன்று பிறழ்ச்சியற்ற மதிப்பீடுகளைக் கொடுக்கும். $\sum_i \sum_j (x_{ij} - \bar{x})^2$ ஐ, $(n - 1)$ ஆல் வகுக்கும்போது மாறுபாட்டளவைக்கான $(n - 1)$ கட்டின்மை எண்ணிக்கையுடன் ஒரு பிறழ்ச்சியற்ற மதிப்பீடாக அமைகிறது. இதுபோலவே 'i' பிரிவில் உள்ள n_i மதிப்புகள் $\bar{x}_i$ கூட்டுச் சராசரியைக் கொண்ட இயைபிலா மாதிரியாதலால் $\sum_i (x_{ij} - \bar{x}_i)^2$ ஐ $n_i - 1$ஆல் வகுக்கும் போதும் $\sum_i \sum_j (x_{ij} - x_i)^2$ ஐ $n - h$ ஆல் வகுக்கும்போதும் மாறுபாட்டளவையின் $(n_i - 1)$, $(n - h)$ கட்டின்மை எண்ணிக்கைகட்கான பிறழ்ச்சியற்ற மதிப்பீடாகும். h என்பது பிரிவுகளின் எண்ணிக்கையைக் குறிக்கும். இவ் விரண்டு மதிப்பீடுகளின் கணக்கியல் எதிர்பார்ப்புகளும் சமமாக அமையவேண்டி உள்ளதால் $\sum n_i (\bar{x}_i - \bar{x})^2$ ஐ $h - 1$ ஆல் வகுக்கும்போது மாறுபாட்டளவைக்கான மற்றொரு பிறழ்ச்சியற்ற மதிப்பீடாகக் கொள்கிறோம் இம் முடிவுகள் யாவையும் அடியிற்காணும் மாறுபாட்டளவை ஆய்வுப்பட்டியல் வடிவில் குறிப்பது வழக்கம்.

மாறுபாட்டளவை ஆய்வுப் பட்டியல்
(Analysis of variance table)

மாறுபாட்டின் மூலம்	கட்டின்மை எண்ணிக்கை	விலக்க வர்க்கங்களின் கூட்டுத்தொகை	மாறுபாட்டளவை	F	சிறப்பெல்லை	
					5%	1%
பிரிவுகளிடையே	$h-1$	$\sum_i n (\bar{x}_i - \bar{x})^2 = D_1$	$\frac{D_1}{h-1} = S_1^2$	$\frac{S_1^2}{S_2^2}$		
பிரிவுகளுக்குள்ளே	$n-h$	$\sum_i \sum_j (x_{ij} - x_i)^2 = D_2$	$\frac{D_2}{n-h} = S_2^2$			
மொத்தம்	$n-1$	$\sum_i \sum_j (x_{ij} - \bar{x})^2$				

மாறுபாட்டளவைகளின் மதிப்பீடுகளுக்கான ஒருமைத்தன்மைக்கான சோதனையாக (test of homogeneity) F-சோதனை செய்கிறோம். பிரிவுகளிடையே ஏற்பட்ட மாறுபாட்டளவைகளுடன் பிரிவுகளுக்குள்ளே ஏற்பட்ட மாறுபாட்டளவையை ஒப்பிடுவது தான் நடைமுறையில் உள்ளது. பிரிவுகளுக்கிடையே அமைந்த மதிப்பீடு, பிரிவுகளுக்குள்ளே அமைந்த மதிப்பீட்டைக் காட்டிலும் குறிப்பிடும்படியான வேறுபாட்டுடன் அமைந்திருந்தால் நாம் பகுப்பதற்காக எடுத்துக்கொண்ட பண்பு மாறியின் அளவுகளில் குறிப்பிடும்படியான விளைவை ஏற்படுத்தி உள்ளது என்ற முடிவை மேற்கொள்ள ஏதுவாய் இருக்கிறது. இவ்விடத்து முழுமைத் தொகுதி ஒருபடித்தானது (homogeneity) என்பது மறுக்கப்பட்டு, பலபடித்தானது (heterogeneous) என்ற முடிவை மேற்கொள்கிறோம். மதிப்பீடுகள் குறிப்பிடும்படியான வித்தியாசத்துடன் அமையாவிட்டால் முழுமைத்தொகுதி ஒருபடித்தானது என்ற முடிவினை மேற்கொள்கிறோம்.

நான்கு இன நெல்லின் விளைபலனில் ஏற்படும் வித்தியாசத்தைக் கணக்கிட ஒரு கட்டளை முறைப்படி அமைக்கும் சோதனையைக் காண்போம். கொடுக்கப்பட்ட நிலத்தைப் பல பாத்திகளாகப் பிரித்து நான்கு இன விதைகளையும் ராண்டமாக விதைப்பதே சோதனை முறையாகும். நான்கு இன விதைகளை முறையே A, B, C, D எனக் கொண்டால், 16 பாத்திகளாகப் பிரிக்கப்பட்ட வயலில் கீழ்க்கண்டவாறு விதைக்கலாம்.

A	*C*	*B*	*A*
C	*A*	*B*	*C*
B	*D*	*D*	*B*
D	*C*	*D*	*A*

விதைகள் கிடைக்கும் தன்மைக்கேற்பவும் நிலத்தின் பரப்பைக் கொண்டும் ஒவ்வோர் இனமும் சமமான எண்ணிக்கையுடைய பாத்திகளிலும் அல்லது A இனம் 4 பாத்திகளிலும், B இனம் 2 பாத்திகளிலுமாகச் சம எண்ணிக்கை இல்லாமலும் விதைக்கலாம். ஒவ்வோர் இனமும் சமமான பாத்திகளில் விதைக்கப்படுவது ஆய்வுமுறையைச் சுலபமாக்கும்.

நிலம் முழுதும் ஒரே தன்மைத்தானது எனக்கொண்டு 16 சமபாகங்களாகப் பிரித்து, ஒவ்வோர் இன விதையும் நான்கு பாத்திகளில் ராண்டமாகப் பயிரிடப்பட்டுள்ளது. (விதைகளை ராண்டமாக விதைக்கப் பின்னிணைப்பில் கொடுக்கப்பட்டுள்ள ராண்டம் எண்களை உபயோகப்படுத்தலாம்.) இதற்காக அமைக்கப்படும் மாறுபாட்டளவை ஆய்வுப்பட்டியலிலிருந்து எடுக்கப்படும் முடிவு, ஒருகட்டளை விதிப்படி அமைக்கப்பட்ட செய்முறைத் திட்டத்திலிருந்து (simple randomised design) எடுக்கப் பட்டதாகக் கொள்ளப்படும்.

விலக்கவர்க்கங்களைக் கணக்கிடும் முறை

நமக்குத் தேவையான விலக்கவர்க்கங்களின் கூட்டுத் தொகையைக் கணக்கிடும்போது, ஒவ்வொன்றிற்குமான கூட்டுச் சராசரியிலிருந்து விலக்கங்களை அமைக்கத் தேவையில்லை. n உறுப்புகள் கொண்ட $(x_1, x_2, \ldots x_n)$ மாதிரியின் மாறுபாட்டளவையை, விலக்கங்கள் அமைக்காமலேயே கணக்கிடும் முறையைக் காண்போம்.

$$\Sigma (x_i - \bar{x})^2 = \Sigma x^2 - n\bar{x}^2 = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}$$

இம் முறையை உபயோகப்படுத்தி இப் பகுதிக்குத் தேவையான விலக்கங்கள் யாவும் கணக்கிடப்படுகின்றன.

$$\sum_i \sum_j (x_{ij} - \bar{x}) = \sum_i \sum_j x_{ij}^2 - T^2/N \qquad (1)$$

$T = \sum_i \sum_j x_{ij}$ மொத்த அளவுகளின் கூட்டுத்தொகை.

$$\Sigma \Sigma (x_{ij} - \bar{x})^2 = \Sigma \Sigma x_{ij}^2 - Ti^2/n_i$$

Ti — i பிரிவில் உள்ள அளவுகளின் கூட்டுத்தொகை.

$$\sum_i \sum_j (x_{ij} \; \bar{x})^2 = \Sigma \Sigma x_{ij}^2 - \Sigma Ti^2/ni \qquad (2)$$

சமன்பாடு (1)-லிருந்து (2) ஐக் கழித்து,

$$\Sigma n_i (\bar{x}_i - \bar{x})^2 = \Sigma \frac{T_i^2}{n_i} - \frac{T^2}{n}$$

எனும் நமக்குத் தேவையான மூன்றாவது விலக்கவர்க்கத்தைப் பெறுகிறோம். இதன் மதிப்பை

$$\Sigma n_i (\bar{x}_i - \bar{x})^2 = \Sigma n_i \bar{x}_i^2 - \frac{(\Sigma n_i \bar{x}_i)^2)}{n}$$

$$= \sum_i \frac{T_i^2}{n_i} - \frac{T^2}{n}$$

என்ற வாய்பாட்டிலிருந்தும் அமைத்துக்கொள்ளலாம்.

காட்டு :

ஒரே தன்மை வாய்ந்த 35 பாத்திகளில் 7 இன நெல் பயிர் செய்யப்படுகிறது. ஒவ்வோர் இனமும் 5 பாத்திகளில் ராண்டமாக விதைத்துக் கிடைத்த விளைச்சலின் முடிவு மரக்காலில் கொடுக்கப்பட்டுள்ளது. நிமிர்வரிசைகள் (column) யாவும் இனங்களைக் குறிக்கின்றன. இனங்களின் விளைச்சலுக்கான வித்தியாசத்தைக் காணும் சோதனையைச் செய்வோம்.

V_1	V_2	V_3	V_4	V_5	V_6	V_7	மொத்தம்
13	15	14	14	17	15	16	104
11	11	10	10	15	9	12	78
10	13	12	15	14	13	13	90
16	18	13	17	19	14	15	112
12	12	11	10	12	10	11	78
62	69	60	66	77	61	67	462

திருத்தக் காரணி (correction factor) = $C.F. = \frac{(462)^2}{35}$

$= 6098{\cdot}4$

மொத்த விலக்கவர்க்கங்களின் கூட்டுத்தொகை = $T.\ S.\ S.$
(total sum of squares)

$$T.\ S.\ S. = 13^2 + 15^2 + \ldots\ldots + 10^2 + 11^2 - C.F.$$

$= 215{\cdot}6$

இனங்களுக்கிடையில் அமைந்த விலக்கங்களின் கூட்டுத்தொகை (Between Varieties) = $B.V.$

$$B.V = \frac{1}{5}[62^2+69^2+60^2+66^2+77^2+61^2+67^2] - C.F.$$

$$= 41·6$$

இனங்களுக்குள் அமைந்த விலக்கவர்க்கங்களின் கூட்டுத்தொகை, (Within Varieties) = $W.V.$

= (மொத்த விலக்க வர்க்கங்களின் கூட்டுத்தொகை) — (இனங்களுக்கிடையில் அமைந்த விலக்கங்களின் கூட்டுத்தொகை)

$$W.V = T.S.S - B.V.$$

$$= 215·6 - 41·6 = 174$$

மாறுபாட்டாய்வுப் பட்டியல் (Anova Table)

மாறுபாட்டு மூலம்	கட்டின்மை எண்ணிக்கை	விலக்க வர்க்கங்களின் கூட்டுத் தொகை	மாறுபாட்டளவை	F	சிறப்பெல்லை 5%	சிறப்பெல்லை 1%
இனங்களுக்கிடையில்	6	41·6	6·933	1 1	2·44	3·53
இனங்களுக்குள்	28	174·0	6·214			
மொத்தம்	34	215·6				

கணக்கிடப்பட்ட மதிப்பு, 0·05 சிறப்பெல்லையில் F-அளவையின் 6, 28 கட்டின்மை எண்ணிக்கைக்கான பட்டியல் மதிப்பைவிடக் குறைவாக இருப்பதால், இனங்களுக்குள் விளைச்சலில் வித்தியாசம் இல்லை எனும் நமது கருதுகோளை ஏற்றுக் கொள்கிறோம்.

ஒவ்வோர் இனமும் சமமான பாத்திகளில் பயிரிடப்பட்ட செய்முறைத் திட்டத்திற்கான ஆய்வினைக் கண்டோம். சில இடங்களில் இனங்கள் யாவும் சம எண்ணிக்கையுள்ள பாத்திகளில் பயிரிடப்படாத சூழல் இருக்கலாம். இவ் விடங்களில்

A இனம் *m* பாத்திகளிலும், *B* இனம் *n* பாத்திகளிலும், *C* இனம் *p* பாத்திகளிலும் பயிரிடப்படலாம் ($m \neq n \neq p$). இதற்கான ஆய்வுமுறையிலும் சம எண்ணிக்கையுள்ள பாத்திகளில் பயிரிடப்பட்ட செய்முறைத் திட்ட ஆய்வுமுறையிலும் அடிப் படைக் கருத்துகளை ஒத்து இருந்தாலும், இனங்களுக்கிடையில் அமையும் விலக்கவர்க்கங்களைக் கணக்கிடும் முறையில் உள்ள சிறு மாற்றத்தை உணர்தல் நலம்.

காட்டு :

A இனம் 27 பாத்திகளிலும், *B* இனம் 15 பாத்திகளிலும் *C* இனம் 12 பாத்திகளிலுமாகப் பயிரிடப்பட்ட செய்முறைத் திட்டத்திற்கான ஆய்வைக் காண்போம்.

A			B		
58	109	87	103	84	88
101	105	94	109	134	106
167	141	112	86	149	64
104	58	98	120	108	82
106	120	65	112	129	22
100	117	82		C	
133	172	133	34	12	20
165	150	116	5	48	32
120	192	117	21	19	24
			20	17	65

A இனத்திற்கான மொத்தம் = 3122

B இனத்திற்கான மொத்தம் = 1496

C இனத்திற்கான மொத்தம் = 317

மொத்தம் = 4935

$$\text{திருத்தக் காரணி} = C.F. = \frac{(4935)^2}{54} = 451004{\cdot}20$$

இனங்களுக்கிடையில் அமைந்த விலக்கவர்க்கங்களின் மொத்தம்

$$= \frac{(3122)^2}{27} + \frac{(1496)^2}{15} + \frac{(317)^2}{12} - C.F.$$

$$= 518570{\cdot}9 - 451004{\cdot}2$$

$$= 67566{\cdot}7$$

மொத்த விலக்கவர்க்கங்களின்
கூட்டுத்தொகை :

$$= (58)^2 + (109)^2 + \ldots\ldots + (17)^2 + (65)^2 - C.F.$$

$$= 564321{\cdot}0 - 45100{\cdot}2$$

$$= 113316{\cdot}8$$

இனங்களுக்குள் அமைந்த விலக்கவர்க்கங்களின் மொத்தம் =

(மொத்த விலக்கவர்க்கங்களின் கூட்டுத்தொகை) — (இனங்களுக்கிடையில் அமையும் விலக்கவர்க்கங்களின் கூட்டுத்தொகை)

$$= 113316{\cdot}8 - 67566{\cdot}7 = 45750{\cdot}1$$

திருத்தக்காரணியும், மொத்த விலக்கவர்க்கங்களின் கூட்டுத் தொகையும் முன்போலவே கணக்கிடப்பட்டுள்ளன. இனங் களுக்கிடையில் உள்ள விலக்கவர்க்கங்களைக் கணக்கிடும்போது ஒவ்வொன்றிற்குமான மொத்த வர்க்கங்களை அவை அமைந்த பாத்திகளின் எண்ணிக்கையால் வகுத்துவரும் ஈவு களின் கூட்டுத்தொகையிலிருந்து திருத்தக்காரணி கழிக்கப் படுகிறது. எல்லா இனங்களும் சமமான பாத்திகளில் அமையும் போது வகுக்கும் எண்ணும் சமமாவதால் இனங்களின் கூட்டுத் தொகையின் வர்க்கங்களைக் கூட்டிப் பின்பு வகுக்கிறோம். சம மற்ற நிலையில், ஒவ்வோர் இனத்தின் கூட்டுத்தொகைக்கான வர்க்கங்கள், அவ் வினம் அமைந்த பாத்திகளின் எண்ணிக்கையால் வகுக்கப்பட்டுள்ளது என்பதை உணர்தல் நலம்.

மாறுபாட்டாய்வுப் பட்டியல் (Anova Table)

மாறுபாட்டு மூலம்	கட்டின்மை எண்ணிக்கை	விலக்க வர்க்கக் கூட்டுத் தொகை	திட்ட விலக்க வர்க்கம்	F	சிறப்பெல்லை	
					5%	1%
இனங்களுக்கிடையில்	2	67566·7	33783·4	37·6	3·18	5·06
இனங்களுக்குள்	51	45750·1	897·1			
மொத்தம்	53	113316·8				

விளைபலனில் இனங்களுக்கிடையில் வித்தியாசம் இல்லை எனும் நமது கருதுகோள் மறுக்கப்பட்டு, விளைபலனில் குறிப் பிடத்தக்க வித்தியாசம் அமைந்துள்ளது என்ற முடிவைச் சொல்கிறோம்.

இருகட்டளை விதிக்கேற்ற பிரிவுகளுக்கான மாறுபாட்டாய்வு முறை (Analysis of variance for two way classification)

n உறுப்புகளின் x_{ij} மதிப்புகள் A, B என்ற இருகட்டளை விதி களுக்குக் கட்டுப்பட்டுப் பிரிக்கப்படுவதாகக் கொள்ளலாம். A என்பது பிரிவுகளையும் (class), B என்பது தொகுப்புகளையும் (group) குறிப்பதாக அமையும். 'h' பிரிவுகளும் 'k' தொகுப்பு களும் உள்ள அமைப்பில் 'hk' உறுப்புகளின் அமைப்பு h படுகைநிலை வரிசைகளைக் (row) கொண்டதாகவும், k நிமிர்நிலை வரிசைகளைக் (column) கொண்டதாகவும் இருக்கும். முன்பு கண்ட 35 பாத்திகள்கொண்ட சோதனையை 7 பாத்திகள் கொண்ட படுகைநிலை வரிசைகளைத் தொகுப்புகளாகவும் (block), 7 இன நெல்லினைப் பிரிவுகளாகவும் (variety) கொண்டால், இனங்கள் நிமிர்நிலை வரிசைகளைக் குறிக்கும். $V_1, V_2 \ldots\ldots V_7$ என்பது இனங்களையும், $B_1, B_2, \ldots\ldots B_5$ என்பது தொகுப்பு களையும் குறித்தால் நமது அமைப்பு அடியிற் கண்டவாறு அமையும்.

	V_1	V_2	V_3	V_4	V_5	V_6	V_7
B_1							
B_2							
B_3							
B_4							
B_5							

ஒவ்வொரு தொகுப்பிலும் (block) நெல்லினங்கள் ஒவ்வொன்றும் ஒரு முறையாவது அமைவது கண்டிப்பாகும்.

ஒருகட்டளை விதிப்படி 16 பாத்திகளில் 4 இன நெல்லைப் பயிரிட அமைத்த ஓர் அமைப்பு

A	*C*	*B*	*A*
C	*A*	*B*	*C*
B	*D*	*D*	*B*
D	*C*	*D*	*A*

என அமையலாம் எனக் குறிப்பிட்டோம். இனங்கள் யாவும் வயல் முழுதும் ராண்டமாகப் பரவி உள்ளதைக் காண்கிறோம். ஒவ்வொரு படுகைநிலை வரிசையையும் ஒரு தொகுப்பாகக் கொண்டால், ஒவ்வோர் இனமும் ஒரு தடவையாவது அமைய வேண்டுமென்ற கட்டளை ஈடுசெய்யப்படாததைக் காணமுடிகிறது. இருகட்டளை விதிப்படி அமையும் செய்முறைத் திட்டங்கள் இக் கட்டளையையும் ஈடு செய்வதாய் இருக்கவேண்டும். 4 இன நெல்லினுக்கான 16 பாத்திகளைக் கொண்ட வயலுக்கு இரு கட்டளை விதிப்படி அமைக்கும் செய்முறைத் திட்டம்.

A	*C*	*D*	*B*
B	*A*	*D*	*C*
D	*B*	*A*	*C*
A	*D*	*C*	*B*

முதல் அமைப்பில் இனங்கள் யாவும் வயல் முழுதும் ராண்டமாகப் பரவி உள்ளதும், இரண்டாவது பிரிவில் இனங்கள் யாவும் ஒவ்வொரு படுகைநிலை வரிசையிலும் ராண்டமாக அமையப் பெற்றிருப்பதும் குறிப்பிடத்தக்கது.

நமக்குத் தேவையான விலக்கவர்க்கங்களின் கூட்டுத்தொகை

$$\sum_i \sum_j (x_{ij} - \bar{x})^2 = \sum_i k (\bar{x}_i - \bar{x})^2 + \sum h (\bar{x}_j - \bar{x})^2$$

$$+ \sum (x_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2 \quad (3)$$

என்ற வாய்பாட்டில் அமைந்த விலக்கவர்க்கங்களை முறையே அவற்றுக்கான கட்டின்மை எண்ணிக்கையால் $(hk - 1)$ $(h - 1)$, $(k - 1)$, $(h - 1)$ $(k - 1)$ வகுத்தால் முழுமைத் தொகுதிக்கான பிறழ்ச்சியற்ற மதிப்பீடுகள் கிடைக்கும். இம் முடிவுகளை மாறுபாட்டளவை ஆய்வுப் பட்டியலில் அமைக்கும் போது அடியிற் கண்டவாறு அமையும்.

மாறுபாட்டாய்வுப் பட்டியல் (Anova table)

மாறுபாட்டு மூலம்	கட்டின்மை எண்ணிக்கை	விலக்க வர்க்கங்களின் கூட்டுத்தொகை	மாறுபாட்டளவை	F	சிறப்பெல்லை 5%	சிறப்பெல்லை 1%
பிரிவுகளுக் கிடையில்	$(h-1)$	$\sum_i k (\bar{x}_i - \bar{x})^2 = D_1$	$\frac{D_1}{h-1} = S_1^2$	$\frac{S_1^2}{S_3^2}$		
தொகுதி களுக் கிடையில்	$(k-1)$	$\sum_j h (\bar{x}_j - \bar{x})^2 = D_2$	$\frac{D_2}{k-1} = S_2^2$	$\frac{S_2^2}{S_3^2}$		
பிழை	$(h-1) \times (k-1)$	$\sum_i \sum_j (x_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2 = D_3$	$\frac{D_3}{(h-1)(k-1)} = S_3^2$			
மொத்தம்		$\sum_i \sum_j (x_{ij} - \bar{x})^2$				

பகுப்பிற்கான பண்பின் மாறுபாடுகளைப் பிரிவுகளுக்கிடையில், தொகுப்புகளுக்கிடையில் என்ற மதிப்பீடுகள் தருவதால் கட்டுப் படுத்த முடியாத மாறுபாட்டளவையின் மதிப்புப் பிழை எனக் குறிக்கப்பட்டுள்ளது.

பிரிவுகளுக்கிடையில், தொகுதிகளுக்கிடையில் அமைந்த மாறுபாட்டளவையின் மதிப்பீடுகள், பிழையினால் கிடைத்த மதிப்பீட்டுடன் ஒப்பிடப்பட்டு முடிவு செய்யப்படுகிறது. பிரிவு களுக்கிடையில், தொகுதிகளுக்கிடையில் கிடைத்த மதிப்புகள் பிழைக்கான மாறுபாட்டளவையைவிடக் குறிப்பீடும் படியான

வித்தியாசத்தைக் கொண்டிருந்தால், முழுமைத் தொகுதியின் ஒரே தன்மைக்கான (homogeneity) கருதுகோள் மறுக்கப்படும்.

காட்டு:

முன்பு கண்ட சோதனையில், இனங்களுக்கான பிரிவுகளுடன் (column - class - variety) தொகுப்புகளுக்கான பகுப்பும் (group-block - row) அமையப்பெற்றிருந்தால், இருகட்டளை விதிப்படி பகுக்கப்பட்ட செய்முறைத் திட்டமாகிவிடும். இவ்விரு பண்புகளும் (இனம், நிலத்தின் வேறுபாடு) விளைவில் உண்டாக்கும் மாறுபாட்டிற்கான சோதனையை அமைப்போம்.

நமக்குத் தேவையான விலக்கவர்க்கங்கள் யாவும், (3) என்ற வாய்பாட்டிலிருந்து பெறலாம். மேற்கண்ட சோதனையில் அமைந்தது போலவே மொத்த மாறுபாட்டளவையும் இனங்களுக்கான மாறுபாட்டளவையும் அமையும். இனங்களுக்குள் அமைந்த மாறுபாட்டளவையை, தொகுப்புகளுக்கானதும், பிழையின்பாற்பட்டதுமாகப் பிரித்தால், இருகட்டளை விதிப்படி பகுக்கப்பட்ட ஆய்விற்கான மாறுபாட்டளவை ஆய்வுப் பட்டியலை அமைக்கலாம். தொகுப்பிற்கானதை மட்டும் கணக்கிட்டு, மொத்தத்திலிருந்து இனம், தொகுப்பு ஆகியவற்றின் விலக்கவர்க்கங்களைக் கழிப்பதால் பிழையின் விலக்கவர்க்கம் கிடைக்கப் பெறுவோம். இவை முறையே 6 (இனம்) 4 (தொகுப்பு), 24 (பிழை), 34 (மொத்தம்) எனக் கட்டின்மை எண்ணிக்கை பெற்றவையாய் அமையும்.

$$\left.\begin{array}{r}\text{தொகுப்புகளுக்கான விலக்க}\\ \text{வர்க்கங்களின் மொத்தம்}\end{array}\right\} = \frac{1}{7}[104^2+78^2+\ldots+78^2]-C.F.$$

$$= 6232{\cdot}57 - 6098{\cdot}4$$

$$= 134{\cdot}12$$

$$\text{பிழை} = 215{\cdot}6 - (41{\cdot}6 + 134{\cdot}12) = 39{\cdot}8$$

மாறுபாட்டாய்வுப் பட்டியல்

மாறுபாட்டினமூலம்	கட்டின்மை எண்ணிக்கை	விலக்கவர்க்கங்களின் மொத்தம்	மாறுபாட்டளவை	F	சிறப்பெல்லை 5%	சிறப்பெல்லை 1%
இனங்களுக்கிடையில்	6	41·6	6·93	4·2	3·84	7·31
தொகுதிகளுக்கிடையில்		134·2	33·55	20·2	5·77	13·93
பிழை	24	39·8	1·66			
மொத்தம்	34	215·6				

இனங்களுக்கான விளைவில் குறிப்பிடும்படியான வித்தியாசம் இல்லை எனும் கருதுகோள் மறுக்கப்படுகிறது. எல்லாத் தொகுப்பு களிலும் நிலவளம் சமமெனும் கருதுகோள் மறுக்கப்படுகிறது. இனங்களுக்கிடையில், தொகுப்புகளுக்கிடையில் அமைந்த F-சோதனையின் அளவு, பட்டியல் மதிப்பைவிட அதிகமாக இருப்பதால் நமது கருதுகோள் மறுக்கப்பட்டு, இனங்களுக்கான விளைபலன் குறிப்பிடும்படியான வித்தியாசத்தைப் பெற்றிருக்கிறது எனவும், தொகுப்புகளுக்கிடையில் நிலவளத்திற்கான தன்மை சமமாக இல்லை எனவும் சொல்கிறோம்.

காட்டு :

பன்றிகளின் எடை உயர்வைக் கணக்கிட டன்லப் (Dunlop) என்பவரால் செய்யப்பட்ட சோதனை முடிவுகளைப் பார்ப்போம். வைட்டமின்கள் மாறுபட்ட A, B, C என்ற மூன்று உணவு முறைகளை நான்கு தொகுப்புகளாக அமைத்த ஓரீற்றுக் குட்டிகளுக்குக் (litter) கொடுத்த சோதனையில் கிடைத்த எடை உயர்வின் முடிவுகள் கொடுக்கப்பட்டுள்ளன. நமது ஆய்விற்கான தொகுப்புகள் யாவும் சோதனை திரும்பச் செய்யப்பட்ட தன்மையைக் குறிக்கும்.

பன்றிகளின் தொகுப்பு

உணவு முறை	I	II	III	IV	
A	7·0	16·0	10·5	13·5	47·0
B	14·0	15·5	15·0	21·0	65·5
C	8·5	16·5	9·5	13·5	48·0
	29·5	48·0	35·0	48·0	160·5

நமக்குத் தேவையான கருதுகோள்கள்,

'உணவுமுறையில் குறிப்பிடும்படியான வித்தியாசம் இல்லை' என்றும், 'தொகுப்பு முறையிலும் குறிப்பிடும்படியான வித்தியாசம் இல்லை' என்றும் அமைகின்றன.

மாறுபாட்டளவை ஆய்வுப் பட்டியலை அமைக்க அடியிற் காணும் முறையில் விலக்கவர்க்கங்களைக் கணக்கிடுகிறோம்.

$$\text{திருத்தக்காரணி} = C.F. = \frac{(160\cdot5)^2}{12} = 2146\cdot69$$

உணவுமுறைகளுக்கிடையில் அமைந்த மாறுபாட்டளவை :

$$= \frac{1}{4}\left[(47)^2 + (65\cdot5)^2 + (48\cdot0)^2\right] - C.F.$$

$$= 2200\cdot81 - 2146\cdot69 = 54\cdot12$$

தொகுப்பு முறைகளுக்கிடையில் அமைந்த மாறுபாட்டளவை :

$$= \frac{1}{3}\left[(29\cdot5)^2 + (48\cdot0)^2 + (35.0)^2 + (48\cdot0)^2\right] - C.F.$$

$$= 2234\cdot42 - 2146\cdot69 = 87\cdot74$$

$$\text{மொத்த விலக்கவர்க்கம்} = \left[(7\cdot0)^2 + (16\cdot0)^2 + \ldots + (13\cdot5)^2\right] - C.F.$$

$$= 2316\cdot75 - 2146\cdot69 = 170\cdot06$$

$$\text{பிழை} = 170\cdot06 - 54\cdot12 - 87\cdot74 = 28\cdot21$$

மாறுபாட்டாய்வுப் பட்டியல்

மாறுபாட்டின் மூலம்	கட்டின்மை எண்ணிக்கை	விலக்க வர்க்கங்களின் மொத்தம்	மாறுபாட்டளவை	F	சிறப்பெல்லை	
					5%	1%
உணவு முறைகளுக்கிடையில்	2	54·12	27·06	5·75	5·14	10·92
தொகுப்பு முறைகளுக்கிடையில்	3	87·73	29·24	6·22	4·76	9·76
பிழை	6	28·21	4·702			
மொத்தம்	11	170·06				

F- சோதனை முடிவுகளின்படி, உணவு முறையிலும் தொகுப்பு முறையிலும் குறிப்பிடும்படியான வித்தியாசம் இல்லை எனும் நமது கருதுகோள் மறுக்கப்பட்டு, இரண்டிலும் குறிப்பிடும்படியான வித்தியாசம் உண்டு என்ற முடிவைச் சொல்கிறோம்.

இரு அளவுகளும் 0·05 சிறப்பெல்லையில் குறிப்பிடும்படியான வித்தியாசத்தைப் பெற்றும், 0·01 சிறப்பெல்லையில் குறிப்பிடும்படியான வித்தியாசத்தைப் பெறுமலும் இருப்பது கவனிக்கத் தக்கது. உணவு முறையில் அமைந்த குறிப்பிடும்படியான வித்தியாசம் அவ்வளவு சிறப்பாக அமைந்ததாகச் சொல்ல முடியாது. தொகுப்பு முறையில் அமைந்த வித்தியாசம் நமது பகுப்பு முறையின் தன்மையைச் சார்ந்ததாகும். ஒவ்வொரு தொகுப்பும் சோதனை திரும்பச் செய்ததற்குச் சமமாகக் கொண்டதால், தொகுப்புகளுக்கிடையில் ஏற்பட்ட சராசரி விலக்கவர்க்கம் உணவு முறையில் ஏற்பட்டதெனச் சொல்ல முடியாது. தொகுப்புமுறை மாறுபட்டிருக்கும்போது, நமது சோதனை உணவு முறைகளுக்குள் மற்றும் உணவுமுறைகளுக் கிடையில் என்பதாய் அமைந்திருக்கும். உணவு முறைகளுக்குள் அமையும் விலக்கவர்க்கம் தற்போதைய அமைப்பைவிட அதிகமாக அமைந்து, செய்முறைத் திட்டத்தின் தன்மையை வலு வற்றதாய்ச் செய்திருக்கும் என்பது குறிப்பிடத்தக்கது.

மூக்கட்டளை விதிப்படி பகுத்த செய்முறைத் திட்டத்தின் மாறு பாட்டளவை ஆய்வுமுறை (Analysis of variance for three way classification)

லேடின் சதுரம்: (Latin Square) மூன்றுகட்டளைப் பகுப்பு முறையில் n, p, k என்பன பகுப்புகளுக்கான எண்ணிக்கைகளைக் குறிப்பதாக அமைந்து $n=p=k$ என்ற கட்டளையை ஏற்கும்போது கிடைக்கும் செய்முறைத் திட்டம் 'லேடின் சதுரம்' எனப்படும். முன்பு கண்ட A, B, C, D என்ற நான்கு இனங்களுக்கான சோதனையை இச் செய்முறைத் திட்டத்தில் அமைக்கும் முறையைக் காண்போம். ஒருகட்டளை விதிப்படி அமைந்த திட்டத்தில் நான்கு இனங்களும் வயல் முழுவதும் (16 பாத்திகளில்) பரவி இருந்ததையும், இரு கட்டளை விதிப்படி அமைந்த செய்முறைத் திட்டத்தில் நான்கு இனங்களும் ஒவ்வொரு படுகைநிலை வரிசையிலும் (row-block) ஒரு தடவையாவது அமைந்திருந்ததையும் கண்டோம். சோதனை முறைக்கேற்ப ஓர் இனத்திற்கான பாத்திகளை அதிகப்படுத்தியும் குறைத்தும் அமைக்கலாம் என்பதையும் கண்டோம். இந்த இரு கட்டளைகளிலும் சிறு மாற்றம் செய்யும்போது 'லேடின் சதுரம்' கிடைக்கப் பெறுகிறோம்.

எல்லா இனங்களும் சம எண்ணிக்கை உள்ள பாத்திகளில் அமைந்து படுகைநிலை வரிசையிலும், நிமிர்நிலை வரிசையிலும் ஒரு தடையாவது அமையப்பெற்றால் 'லேடின் சதுரம்' கிடைக்கும்.

A	*C*	*D*	*B*
B	*D*	*A*	*C*
D	*B*	*C*	*A*
C	*A*	*B*	*D*

முக்கட்டளை விதிப் பகுப்பு முறையின் லேடின் சதுர ஆய்வு முறையை அடியிற்காணும் சோதனையில் காணலாம். நமது ஆய்விற்கான மாறி ஓர் ஏக்கருக்கான விளைபலனாக இருக்கட்டும். நமது சோதனையின் நோக்கம் விளைவில் ஏற்பட்ட மாற்றம்

(1) இனங்களால் ஏற்பட்டதா
(2) உரத்தால் ஏற்பட்டதா
(3) விளைநிலத்தால் ஏற்பட்டதா

என்பதை அறிவதாக இருக்கலாம்

விளைநிலத்தை 'n^2' பாத்திகளாக அமைத்து, n-நிமிர்நிலை வரிசைகளாகவும், n-படுகைநிலை வரிசைகளாகவும் அமைக்கிறோம். 'n' இனங்களும் ராண்டமாக விதைக்கப்படுகின்றன. ஒவ்வோர் இனமும் நிமிர்நிலை வரிசையிலும் படுகைநிலை வரிசையிலும் ஒரு தடவையாவது அமையும்படி விதைக்கப்படும்போது லேடின் சதுரம் (Latin square) கிடைக்கப் பெறுகிறோம். மூன்று இன (A, B, C) நெல்லினைப் பயிரிட்டு விளைபலனைக் காண D_0, D_1, D_2 என்ற மூன்று உரங்கள் கொடுக்கப்படுகின்றன. கொடுக்கப்பட்ட நிலத்தை '3^2' பாத்திகளாகப் பிரித்து A, B, C என்ற மூன்று இனமும் நிமிர், படுகைநிலை வரிசைகளில் அமையும்படி விதைக்கப்படுகின்றன.

A	*C*	*B*
C	*B*	*A*
B	*A*	*C*

குறிப்பிட்ட இன எண்ணிக்கைகளுக்கான 'லேடின் சதுர' அமைப்புகள் புள்ளிவிவரப் பட்டியல்களில் அமையப்பெற்றிருக்கும். ராண்டமாக அமையப்பெற்ற இவ் வமைப்புகளை உபயோகப் படுத்துவது தவற்றினைக் குறைக்க ஏதுவாய் இருக்கும்.

நமத்குத் தேவையான விலக்க வர்க்கங்கள் யாவும்

$$\Sigma\Sigma (x_{ij} - \bar{x})^2 = n \sum_i (\bar{x}_i - \bar{x})^2 + n \sum_j (\bar{x}_j - \bar{x})^2 + n \sum_k (\bar{x}_k - \bar{x})^2 + \sum_i \sum_j (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j - \bar{x}_k + 2\bar{x})^2$$

என்ற வாய்பாட்டில் அமைந்துள்ளன. இவற்றுக்கான கட்டின்மை எண்ணிக்கைகள் முறையே

$(n^2 - 1)$, $(n - 1)$, $(n - 1)$, $(n - 1)$, $(n - 1)(n - 2)$ என அமைகின்றன.

மாறுபாட்டாய்வுப் பட்டியல்

மாறுபாட்டின் மூலம்	கட்டின்மை எண்ணிக்கை	விலக்க வர்க்கங்களின் மொத்தம்	விலக்க வர்க்கம்	F	சிறப்பெல்லை 5%	சிறப்பெல்லை 1%
இனங்களுக்கிடையில்	$(n-1)$	$\Sigma (\bar{x}_i - \bar{x})^2 = D_1$	$\frac{D_1}{n-1} = S_1^2$	$\frac{S_1^2}{S_4^2}$		
படுகை நிலைவரிசைகளுக்கிடையில்	$(n-1)$	$\Sigma (\bar{x}_j - \bar{x})^2 = D_2$	$\frac{D_2}{n-1} = S_2^2$	$\frac{S_2^2}{S_4^2}$		
நிமிர் நிலை வரிசைகளுக்கிடையில்	$(n-1)$	$\Sigma (\bar{x}_k - \bar{x}) = D_3$	$\frac{D_3}{n-1} = S_3^2$	$\frac{S_3^2}{S_4^2}$		
பிழை	$(n-1) \times (n-2)$	$\Sigma\Sigma (x_{ij} - \bar{x}_i - \bar{x}_j - \bar{x}_k + 2\bar{x})^2 = D_4$	$\frac{D_4}{(n-1)(n-2)} = S_4^2$			
மொத்தம்	(n^2-1)	$\Sigma\Sigma (x_{ij} - \bar{x})^2$				

காட்டு ;

கோதுமையின் விளைச்சலைக் காணச் செய்யப்பட்ட லேடின் சதுரச் செய்முறைத் திட்ட முடிவுகள் தரப்பட்டுள்ளன. *A*, *B*,

C, D, E என்பன இனங்களேயும், நிமிர்நிலே வரிசைத் தொகுப்பையும், படுகைநிலே வரிசை உரத்தன்மையையும் குறிப்பதாகக் கொண்டு, விளைபலனில் ஏற்பட்ட மாற்றத்திற்கான தன்மையைக் கணக்கிடுவோம்.

A	16	B	10	D	9	E	9	C	11
E	10	C	9	B	12	D	11	A	14
B	15	D	8	C	10	A	18	E	18
D	12	E	6	A	13	C	12	B	13
C	13	A	11	E	7	B	14	D	10

இனங்களுக்கான மொத்தம்: $A = 72$, $B = 64$, $C = 55$, $D = 50$, $E = 50$

படுகைநிலே வரிசைகளின் கூட்டுத்தொகை : $R_1 = 55$, $R_2 = 56$, $R_3 = 69$, $R_4 = 56$, $R_5 = 55$

நிமிர்நிலே வரிசைகளின் கூட்டுத்தொகை: $C_1 = 66$, $C_2 = 44$, $C_3 = 51$, $C_4 = 64$, $C_5 = 66$

மொத்தம் $= 291$

$$\text{திருத்தக்காரணி} = C.F. = \frac{(291)^2}{25} = 3387 \cdot 24$$

$$\left.\begin{array}{r}\text{இனங்களுக்கான விலக்க}\\\text{வர்க்கம்}\end{array}\right\} = \frac{1}{5}[(72)^2 + (64)^2 + (55)^2 + (50)^2 + (50)^2] - C.F.$$

$$= 3461 \cdot 0 - 3387 \cdot 24$$

$$= 73 \cdot 76$$

படுகைநிலை வரிசைகளுக்கான விலக்கவர்க்கம் $= \frac{1}{5} [(55)^2 + (56)^2 + (69)^2 + (56)^2 + (55)^2] - C.F.$

$= 3416 \cdot 6 - 3387 \cdot 24$

$= 29 \cdot 36$

நிமிர்நிலை வரிசைகளுக்கான விலக்கவர்க்கம் $= \frac{1}{5} [(66)^2 + (44)^2 + (51)^2 + (64)^2 + (66)^2] - C.F.$

$= 3469 \cdot 00 - 3387 \cdot 24$

$= 81 \cdot 76$

மொத்த விலக்க வர்க்கங்களின் கூட்டுத்தொகை $= (16)^2 + (10)^2 + \ldots + (14)^2 + (10)^2 - C.F$

$= 3615 \cdot 00 - 3387 \cdot 24$

$= 227 \cdot 76$

பிழையின்பாற்பட்ட விலக்கவர்க்கங்களின் மொத்தம் = (மொத்த விலக்கவர்க்கம்) (இனம், நிமிர்நிலை வரிசை, படுகைநிலை — வரிசை ஆகியவற்றிற்கான விலக்க வர்க்கங்களின் கூட்டுத்தொகை)

$= 227 \cdot 76 - (73 \cdot 76 + 29 \cdot 36 + 81 \cdot 76) = 42 \cdot 88$

மாறுபாட்டாய்வுப் பட்டியல்

மாறுபாட்டின் மூலம்	கட்டின்மை எண்ணிக்கை	விலக்க வர்க்கங்களின் மொத்தம்	விலக்க வர்க்கம்	F	சிறப்பெல்லை	
இனங்களுக்கிடையில்	4	73·76	18·44	5·2	3·11	5·06
படுகைநிலை வரிசைகளுக்கிடையில்	4	29·36	7·34	2·06	3·11	5·06
நிமிர்நிலை வரிசைகளுக்கிடையில்	4	81·76	20·44	5·73	3·11	5·06
பிழை	12	42·88	3·57			
மொத்தம்	24	227·76				

இனங்களுக்கிடையிலும், நிமிர்நிலை வரிசைகளுக்கிடையிலும் குறிப்பிடும்படியான வித்தியாசம் இல்லை எனும் கருதுகோள்கள் மறுக்கப்பட்டும், படுகைநிலை வரிசைகளுக்கிடையில் வித்தியாசம் இல்லை எனும் கருதுகோள் ஏற்கப்பட்டும் முடிவு செய்யப் படுகிறது. படுகைநிலை வரிசைகள் உரங்களை (D_1, D_2, D_3, D_4, D_5)-க் குறிக்கும். இனங்களிலும் (A, B, C, D, E), தொகுப்பு முறையிலும் (C_1, C_2, C_3, C_4, C_5,) குறிப்பிடும்படியான வித்தியாசம் உண்டு என்பதைச் சோதனை முடிவு உறுதிப்படுத்துகிறது. உரத்தின் தன்மை விளைபலனில் குறிப்பிடும்படியான வித்தியாசத்தை உண்டாக்கியதாகச் சொல்ல இடமில்லை என்ற முடிவையும் சொல்கிறோம்.

பயிற்சி :

1. 4 தொகுதிகளாகப் பிரிக்கப்பட்ட விளக்குகளின் எரியும் காலம் மணியில் கொடுக்கப்பட்டுள்ளது. நான்கு தொகுதி களிலும் அமைந்த விளக்குகளின் எரியும் காலத்திற்கான கூட்டுச் சராசரியில் குறிப்பிடத்தக்க வித்தியாசம் உண்டா என்பதற்கான சோதனையைச் செய்.

B_1	1600	1610	1650	1680	1670	1720	1800	
B_2	1580	1640	1640	1700	1750			
B_3	1460	1550	1600	1620	1640	1660	1740	1820
B_4	1510	1520	1530	1570	1600	1680		

2. 20 பன்றிகளை 5 பன்றிகள் கொண்ட 4 தொகுதி களாகப் பிரித்து, ஒவ்வொரு தொகுதிக்கும் ஒருவிதமான உணவு கொடுக்கப்பட்டது. குறிப்பிட்ட காலத்திற்குள் ஏற்பட்ட பன்றிகளின் எடை உயர்வு கொடுக்கப்பட்டுள்ளது. எடை உயர்விற்கான தன்மையில் நான்கு உணவுகளிலும் குறிப்பிடும் படியான வித்தியாசம் உண்டா என்பதைச் சோதனை செய்.

உணவு A	உணவு B	உணவு C	உணவு D
183	168	210	196
144	148	233	184
135	152	220	199
140	146	226	187
143	157	229	193

3. ஐந்து ஆய்வாளர்கள், ஐந்து நகரங்களுக்குத் தனித்தனியாகக் கணக்கிட்ட வாழ்க்கைக் குறியீட்டெண்கள் (Cost of living index) கொடுக்கப்பட்டுள்ளன.

நகரம்	ஆய்வாளர்கள்				
	1	2	3	4	5
1	270	263	264	263	260
2	280	265	274	274	279
3	275	284	278	271	296
4	271	269	272	297	274
5	279	267	269	263	284

இப் பட்டியலுக்கான 'மாறுபாட்டாய்வுப் பட்டியலை அமைத்து, ஆய்வாளர்களுக்கிடையிலும், நகரங்களுக்கிடையிலும்,

(வாழ்க்கைக் குறியீட்டெண்ணில்) குறிப்பிடும்படியான வித்தியாசம் உண்டா என்பதைக் காண்.

4. ஆறு டின்களில் உள்ள மருந்துப் பொடியில் உள்ள ஈரத் (moisture) தன்மையைக் காண, நான்கு நபர்கள் ஒவ்வொரு டின்னிலிருந்தும் ஒரு மாதிரி எடுத்துச் சோதனை செய்த முடிவுகள் கொடுக்கப்பட்டுள்ளன.

ஆய்வாளர்கள்	டின்கள்					
	1	2	3	4	5	6
1	9	10	9	10	11	11
2	12	11	9	11	10	10
3	11	10	10	12	11	10
4	12	13	11	14	12	10

இவ் அளவுகளுக்கான மாறுபாட்டாய்வுப் பட்டியலை அமைத்து, சோதனையாளர்களுக்கிடையிலும், டின்களுக்கிடையிலும் உள்ள வேறுபாட்டின் தன்மைபற்றி விவரி.

5. 9 இனக் கருப்புகள் ஒவ்வொன்றையும் 5 பாத்திகளில் பயிரிட்டு ஓர் ஏக்கருக்கான விளைபலன் டன்னில் கொடுக்கப்பட்டுள்ளது. செய்முறைத் திட்டம், 5 தொகுதிகளிலும் ராண்டம் முறையில் பயிரிடப்பட்டதெனக் கொண்டு இனங்களுக்கான விளைபலனில் உள்ள வித்தியாசத்திற்கான சோதனையைச் செய். அதிக விளைச்சல் தந்துள்ள இரண்டு இனங்களின் விளைபலனுக்கான வித்தியாசம் குறிப்பிடும்படியாக உள்ளதா என்பதையும் சோதனை செய்.

இனம் \ தொகுதி	B_1	B_2	B_3	B_4	B_5
1	1·94	2·08	1·86	2·21	2·03
2	1·70	1·96	1·83	1·60	2·13
3	2·23	2·26	2·23	2·08	2·02
4	2·14	2·08	2·16	2·16	2·17
5	1·80	2·23	1·67	2·11	2·01
6	1·82	2·06	2·03	1·96	2·28
7	1·91	2·06	2·22	2·14	2·28
8	1·90	2·25	1·92	1·99	2·02
9	1·98	2·03	1·81	1·77	1·88

6. 5 வித உணவு முறைகளை, 5 ஓரீற்றுக் குட்டிகள் அடங்கிய 8 தொகுப்புகளுக்குக் கொடுத்து, எடை உயர்விற்கான மதிப்புக் குறிக்கப்பட்டுள்ளது. தொகுப்பு முறையும், உணவின் தன்மையும் எடை உயர்வில் உள்ள நிலையைக் காண்.

தொகுதி \ இனம்	1	2	3	4	5
1	57·0	64·8	70·7	68·3	76·0
2	55·0	66·6	59·4	67·1	74·5
3	62·1	69·5	64·5	69·5	76·5
4	74·5	61·1	74·0	72·3	86·6
5	86·7	91·8	78·5	90·6	94·7
6	42·0	51·8	55·8	44·3	43·2
7	71·9	69·2	63·0	53·8	61·1
8	51·5	48·6	48·1	40·9	54·4

7. உருளைக் கிழங்கு விளைச்சலுக்கான லேடின் சதுர (Latin Square) செய்முறைத் திட்டத்தின் முடிவு கொடுக்கப்பட்டுள்ளது.

O — கட்டுப்படுத்தப்பட்டது (உரமிடப்படாதது)
C — சினோமைடு (Sinomide)
CC — சினோமைடு இரண்டு பங்கு
S — அமோனியாக் கந்தகி (Sulphate of amonia)
SS — அமோனியாக் கந்தகி இரண்டு பங்கு

130	CC	120	S	43	O	118	C	137	SS
70	S	81	C	84	SS	104	CC	72	O
120	SS	109	CC	82	C	180	O	71	S
42	O	97	SS	97	CC	72	S	48	C
48	C	32	O	71	S	72	SS	45	CC

இச் சோதனைக்கான மாறுபாட்டாய்வுப் பட்டியலை அமை.

8. கோதுமை விளைச்சலுக்கான லேடின் சதுர செய்முறைத் திட்ட முடிவுகள் குறிக்கப்பட்டுள்ளன. லேடின் சதுர மாறுபாட்டாய்வுப் பட்டியலை அமை.

B	4·9	D	6·4	E	3·3	A	9·5	C	11·8
C	9·3	A	4·0	B	6·2	E	5·1	D	5·4
D	7·6	C	15·4	A	6·5	B	6·0	E	4·6
E	5·3	B	7·6	C	13·2	D	8·6	A	4·9
A	9·3	E	6·3	D	11·8	C	15·9	B	7·6

9. *A*, *B*, *C*, D எனும் நான்கு உர வகைகள் விளைபலனில் ஏற்படுத்தும் வித்தியாசத்தைக் காணும் செய்முறைத் திட்டத்தில் 4×4 லேடின் சதுரம் உபயோகப்படுத்தப்பட்டது. 4 உரங்களுக்குமான சராசரி விளைபலன் சமம் எனும் கருதுகோளைச் சோதனை செய்.

A	17	*B*	13	*C*	19	D	15
B	13	*C*	20	D	14	*A*	16
C	21	D	16	*A*	15	*B*	15
D	14	*A*	16	*B*	14	*C*	20

10. ஆறு இன விதைகளின் (*A, B, C, D, E, F*) விளைச்சல் தரம் காணச் செய்யப்பட்ட 'லேடின் சதுர' சோதனை முடிவுகள் தரப்பட்டுள்ளன. இதற்கான மாறுபாட்டாய்வுப் பட்டியலை அமை. *A*, D இன விதைகளின் விளைவிற்கான வித்தியாசம் குறிப்பிடும்படி (significant) உள்ளதா என்பதைக் கணக்கிடு.

B 98	*E* 111	D 113	*C* 118	*A* 60	*F* 130
D 112	*B* 89	*E* 115	*A* 80	*F* 96	*C* 130
A 78	D 106	*B* 100	*F* 105	*C* 103	*E* 118
D 109	*A* 78	*C* 106	*F* 112	*E* 103	*B* 98
E 105	*C* 107	*F* 103	*B* 96	D 112	*A* 72
C 102	*F* 98	*A* 70	*E* 104	*B* 93	D 120

TABLE 1

Table of Ordinates and Area of the Normal curve

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

z	$\phi(z)$	$\int_z^\infty \phi(z)\,dz$	z	$\phi(z)$	$\int_z^\infty \phi(z)\,dz$
·00	·3989	·50000	.25	·3867	·40129
·01	·3989	·49601	·26	·3857	·39743
·02	·3989	·49202	·27	·3847	·39358
·03	·3988	·48803	·28	·3836	·38974
·04	·3986	·48405	·29	·3825	·38591
·05	·3984	·48006	·30	·3814	·38209
·06	·3982	·47608	·31	·3802	·37828
·07	·3980	·47210	·32	·3790	·37448
·08	·3977	·46812	·33	·3778	·37070
·09	·3973	·46414	·34	·3765	·36693
·10	·3970	·46017	·35	·3752	·36317
·11	·3965	.45620	·36	·3739	·35942
·12	·3961	·45224	·37	·3725	·35569
·13	·3956	·44828	·38	·3712	·35197
·14	·3951	·44433	·39	·3697	·34827
·15	·3945	·44038	·40	·3683	·34458
·16	·3939	43644	·41	·3668	·34090
·17	·3932	·43251	·42	·3653	·33724
·18	·3925	·42858	·43	·3637	·33360
·19	·3918	·42465	·44	·3621	·32997
·20	·3910	·42074	·45	·3605	·32636
·21	·3902	·41683	·46	3589	·32276
·22	·3894	·41294	·47	3572	·31918
·23	·3885	·40905	·48	·3555	·31561
·24	·3876	·40517	·49	·3538	·31207

z	$\phi(z)$	$\int_z^\infty \phi(z)\,dz$	z	$\phi(z)$	$\int_z^\infty \phi(z)\,dz$
·50	·3521	·30854	·85	·2780	·19766
·51	·3503	·30503	·86	·2756	·19489
·52	·3485	·30153	·87	·2732	·19215
·53	·3467	·29806	·88	·2709	·18943
·54	·3448	·29460	·89	·2685	·18673
·55	·3429	·29116	·90	·2661	·18406
·56	·3410	·28774	·91	·2637	·18141
·57	·3391	·28434	·92	·2613	·17879
·58	·3372	·28096	·93	·2589	·17619
·59	·3352	·27760	·94	·2565	·17361
·60	·3332	·27425	·95	·2541	·17106
·61	·3312	·27093	·96	·2516	·16853
·62	·3292	·26763	·97	·2492	·16602
·63	·3271	·26435	·98	·2468	·16354
·64	·3251	·26109	·99	·2444	·16109
·65	·3230	·25785	1·00	·2420	·15866
·66	·3209	·25463	1·01	·2396	·15625
·67	·3187	·25143	1·02	·2371	·15386
·68	·3166	·24825	1·03	·2347	·15151
·69	·3144	·24510	1·04	·2323	·14917
·70	·3123	·24196	1·05	·2299	·14686
·71	·3101	·23885	1·06	·2275	·14457
·72	·3079	·23576	1·07	·2251	·14231
·73	·3056	·23270	1·08	·2227	·14007
·74	·3034	·22965	1·09	·2203	·13786
·75	·3011	·22663	1·10	·2179	·13567
·76	·2989	·22363	1·11	·2155	·13350
·77	·2966	·22065	1·12	·2131	·13136
·78	·2943	·21770	1·13	·2107	·12924
·79	·2920	·21476	1·14	·2083	·12714
·80	·2897	·21186	1·15	·2059	·12507
·81	·2874	·20897	1·16	·2036	·12302
·82	·2850	·20611	1·17	·2012	·12100
·83	·2827	·20327	1·18	·1989	·11900
·84	·2803	·20045	1·19	·1965	·11702

z	$\phi(z)$	$\int_z^{\infty} \phi(z)\,dz$	z	$\phi(z)$	$\int_z^{\infty} \phi(z)\,dz$
1·20	·1942	·11507	1·55	·1200	·060571
1·21	·1919	·11314	1·56	·1182	·059380
1·22	·1895	·11123	1·57	·1163	·058208
1·23	·1872	·10935	1·58	·1145	·057053
1·24	·1849	·10749	1·59	·1127	·055917
1·25	·1826	·10565	1·60	·1109	·054799
1·26	·1804	·10383	1·61	·1092	·053699
1·27	·1781	·10204	1·62	·1074	·052616
1·28	·1758	·10027	1·63	·1057	·051551
1·29	·1736	·098525	1·64	·1040	·050503
1·30	·1714	·096800	1·65	·1023	·049471
1·31	·1691	·095098	1·66	·1006	·048457
1·32	·1669	·093418	1·67	·0989	·047460
1·33	·1647	·091759	1·68	·0973	·046479
1·34	·1626	·090123	1·69	·0957	·045514
1·35	·1604	·088508	1·70	·0940	·044565
1·36	·1582	·086915	1·71	·0925	·043633
1·37	·1561	·085343	1·72	·0909	·042716
1·38	·1539	·083793	1·73	·0893	·041815
1·39	·1518	·082264	1·74	·0878	·040930
1·40	·1497	·080757	1·75	·0863	·040059
1·41	·1476	·089270	1·76	·0848	·039204
1·42	·1456	·077804	1·77	·0833	·038364
1·43	·1435	·076359	1·78	·0818	·037538
1·44	·1415	·074934	1·79	·0804	·036727
1·45	·1394	·07352	1·80	·0790	·035930
1·46	·1374	·072145	1·81	·0775	·035148
1·47	·1354	·070781	1·82	·0761	·034380
1·48	·1334	·069437	1·83	·0748	·033625
1·49	·1315	·068112	1·84	·0734	·032884
1·50	·1295	·066807	1·85	·0721	·032157
1·51	·1276	·065522	1·86	·0707	·031443
1·52	·1257	·064255	1·87	·0694	·030742
1·53	·1238	·063008	1·88	·0681	·030054
1·54	·1219	·061780	1·89	·0669	·029379

z	$\phi(z)$	$\int_z^\infty \phi(z)\,dz$	z	$\phi(z)$	$\int_z^\infty \phi(z)\,dz$
1·90	·0656	·028717	2·25	·0317	·012224
1·91	·0644	·028067	2·26	·0310	·011911
1·92	·0632	·027429	2·27	·0303	·011604
1·93	·0620	·026803	2·28	·0297	·011304
1·94	·0608	·026190	2·29	·0290	·011011
1·95	·0596	·025588	2·30	·0283	·010724
1·96	·0584	·024998	2·31	·0277	·010444
1·97	·0573	·024419	2·32	·0270	·010170
1·98	·0562	·023852	2·33	·0264	·0099031
1·99	·0551	·023295	2·34	·0258	·0096419
2·00	·0540	·022750	2·35	·0252	·0093867
2·01	·0529	·022216	2·36	·0246	·0091375
2·02	·0519	·021692	2·37	·0241	·0088940
2·03	·0508	·021178	2·38	·0235	·0086563
2·04	·0498	·020675	2·39	·0229	·0084242
2·05	·0488	·020182	2·40	·0224	·0081975
2·06	·0478	·019699	2·41	·0219	·0079763
2·07	·0468	·019226	2·42	·0213	·0077603
2·08	·0459	·018763	2·43	·0208	·0075494
2·09	·0449	·018309	2·44	·0203	·0073436
2·10	·0440	·017864	2·45	·0198	·0071428
2·11	·0431	·017429	2·46	·0194	·0069469
2·12	·0422	·017003	2·47	·0189	·0064556
2·13	·0413	·016586	2·48	·0184	·0065691
2·14	·0404	·016177	2·49	·0180	·0063842
2·15	·0396	·015778	2·50	·0175	·0052097
2·16	·0387	·015386	2·51	·0171	·0060366
2·17	·0379	·015003	2·52	·0167	·0058677
2·18	·0371	·014629	2·53	·0163	·0057031
2·19	·0363	·014262	2·54	·0158	·0055426
2·20	·0355	·013903	2·55	·0154	·0053861
2·21	·0347	·013553	2·56	·0151	·0052336
2·22	·0339	·013209	2·57	·0147	·0050849
2·23	·0332	·012874	2·58	·0143	·0049400
2.24	·0325	·012545	2·59	·0139	·0047988

z	$\phi(z)$	$\int_z^\infty \phi(z)\,dz$	z	$\phi(z)$	$\int_z^\infty \phi(z)\,dz$
2·60	·0136	·0046612	2·85	·0069	·0021860
2·61	·0132	·0045271	2·86	·0067	·0021182
2·62	·0129	·0043965	2·87	·0065	·0020524
2·63	·0126	·0042692	2·88	·0063	·0019884
2·64	·0122	·0041453	2·89	·0061	·0019262
2·65	·0119	·0040246	2·90	·0060	·0018658
2·66	·0116	·0039070	2·91	·0058	·0018710
2·67	·0113	·0037926	2·92	·0056	·0017502
2·68	·0110	·0036811	2·93	·0055	·0016948
2·69	·0107	·0035726	2·94	·0053	·0016411
2·70	·0104	·0034670	2·95	·0051	·0015889
2·71	·0101	·0033642	2·96	·0050	·0015382
2·72	·0099	·0032641	2·97	·0048	·0014890
2·73	·0096	·0031667	2·98	·0047	·0014412
2·74	·0093	·0030720	2·99	·0046	·0013949
2·75	·0091	·0029798	3·00	·0044	·00134990
2·76	·0088	·0028901	3·10	·0033	·00096760
2·77	·0086	·0028028	3·20	·0024	·00068714
2·78	·0084	·0027179	3·30	·0017	·00048342
2·79	·0081	·0026354	3·40	·0012	·00033693
2·80	·0079	·0025551	3·50	·0009	·00023263
2·81	·0077	·0024771	3·60	·0006	·00015911
2·82	·0075	·0024012	3·70	·0004	·00010780
2·83	·0073	·0023274	3·80	·0003	·00007235
2·84	·0071	·0022557	3·90	·0002	·00004809

Appendix table 1 is taken from tables II, II_1 of Fisher and Yates 'Statistical tables for Biological, Agricultural and Medical Research,' published by Oliver & Boyd, Edinburgh, and by permission of the authors and publishers.

TABLE 2

Table of t

Degrees of freedom (கட்டின்மை எண்ணிக்கை)	Probability of a deviation greater than *t*	
	1%	5%
1	63·657	15·706
2	9·925	4·303
3	5 841	3·182
4	4·604	2·776
5	4·032	2·571
6	3·707	2·447
7	3·499	2·365
8	3·355	2·306
9	3·250	2·262
10	3·169	2·228
11	3·106	2·201
12	3·055	2·179
13	3·012	2·160
14	2·977	2·145
15	2·947	2·131
16	2·921	2·120
17	2·898	2·110
18	2·878	2·101
19	2·861	2·093
20	2·845	2·086
21	2·831	2·080
22	2·819	2·074
23	2·807	2·069
24	2·797	2·064
25	2·787	2·060

Degrees of freedom கட்டின்மை எண்ணிக்கை	Probability of a deviation greater than t	
	1%	5%
26	2·779	2·056
27	2·771	2·052
28	2·763	2·048
29	2·756	2·045
30	2·750	2·042
40	2·704	2·021
60	2·660	2·000
80	2·639	1·990
100	2·626	1·985
∞	2·576	1·960

This table is extracted from table 4·1 of C. R. Rao, S. K. Mitra and A. Mathai—'Formulae and tables for Statistical work'—Statistica Publishing Society, Calcutta—with permission of the authors and Publishers.

TABLE 3

***Table of* chi—square**

Degrees of freedom (கட்டின்மை எண்ணிக்கை)	Probability of a deviation greater than χ^2	
	1%	5%
1	6·63	3·84
2	9·21	5·99
3	11·34	7·81
4	13·28	9·49
5	15·09	11·07
6	16·81	12·59
7	18·48	14·07
8	20·09	15·51
9	21·67	16·92
10	23·21	18·31
11	24·72	19·68
12	26·22	21·03
13	27·69	22·36
14	29·14	23·68
15	30·58	25·00
16	32·00	26·30
17	33·41	27·59
18	34·81	28·87
19	36·19	30·14
20	37·57	31·41
21	38·93	32·67
22	40·29	33·92
23	41·64	35·17
24	42·98	36·42
25	44·31	37·65

Degrees of freedom கட்டின்மை எண்ணிக்கை	Probability of a deviation greater than X^2	
	1%	5%
26	45·64	38·89
27	46·96	40·11
28	48·28	41·34
29	49·59	42·56
30	50·89	43·77
35	57·34	49·80
40	63·69	55·76
50	76·15	67·51
60	88·38	79·08
70	100·42	90·53
80	112·33	101·88
90	124·12	113·15
100	135·81	124·34

This table is extracted from table 5·1 of C. R. Rao, S. K. Mitra and A. Mathai—'Formulae and tables for statistical work'—Statistical Publishing Society, Calcutta—with the permission of the authors and Publishers.

APPENDIX TABLE 4

5% (Roman type) and 1% (Bold face type) Points for the Distribution of F

கட்டின்மை எண்ணிக்கை

f_1 Degrees of Freedom (for greater mean square)

f_2	1	2	3	4	5	6	7	8	9	10	11	12
1	161	200	216	225	230	234	237	239	241	242	243	244
	4,052	**4,999**	**5,403**	**5,625**	**5,764**	**5,859**	**5,928**	**5,981**	**6,022**	**6,056**	**6,082**	**6,106**
2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41
	98.49	**99.01**	**99.17**	**99.25**	**99.30**	**99.33**	**99.34**	**99.36**	**99.38**	**99.40**	**99.41**	**99.42**
3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.76	8.74
	34.12	**30.81**	**29.46**	**28.71**	**28.24**	**27.91**	**27.67**	**27.49**	**27.34**	**27.23**	**27.13**	**27.05**
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91
	21.20	**18.00**	**16.69**	**15.98**	**15.52**	**15.21**	**14.98**	**14.80**	**14.66**	**14.54**	**14.45**	**14.37**
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68
	16.26	**13.27**	**12.06**	**11.39**	**10.97**	**10.67**	**10.45**	**10.27**	**10.15**	**10.05**	**9.96**	**9.89**
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00
	13.74	**10.92**	**9.78**	**9.15**	**8.75**	**8.47**	**8.26**	**8.10**	**7.98**	**7.87**	**7.79**	**7.72**
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57
	12.25	**9.55**	**8.45**	**7.85**	**7.46**	**7.19**	**7.00**	**6.84**	**6.71**	**6.62**	**6.54**	**6.47**
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28
	11.26	**8.65**	**7.59**	**7.01**	**6.63**	**6.37**	**6.19**	**6.03**	**5.91**	**5.82**	**5.74**	**5.67**
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07
	10.56	**8.02**	**6.99**	**6.42**	**6.06**	**5.80**	**5.62**	**5.47**	**5.35**	**5.26**	**5.18**	**5.11**
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91
	10.04	**7.56**	**6.55**	**5.99**	**5.64**	**5.39**	**5.21**	**5.06**	**4.95**	**4.85**	**4.78**	**4.71**
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79
	9.65	**7.20**	**6.22**	**5.67**	**5.32**	**5.07**	**4.88**	**4.74**	**4.63**	**4.54**	**4.46**	**4.40**
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69
	9.33	**6.93**	**5.95**	**5.41**	**5.06**	**4.82**	**4.65**	**4.50**	**4.39**	**4.30**	**4.22**	**4.16**
13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.63	2.60
	9.07	**6.70**	**5.74**	**5.20**	**4.86**	**4.62**	**4.44**	**4.30**	**4.19**	**4.10**	**4.02**	**3.96**
14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.56	2.53
	8.86	**6.51**	**5.56**	**5.03**	**4.69**	**4.46**	**4.28**	**4.14**	**4.03**	**3.94**	**3.86**	**3.80**
15	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55	2.51	2.48
	8.68	**6.36**	**5.42**	**4.89**	**4.56**	**4.32**	**4.14**	**4.00**	**3.89**	**3.80**	**3.73**	**3.67**
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.42
	8.53	**6.23**	**5.29**	**4.77**	**4.44**	**4.20**	**4.03**	**3.89**	**3.78**	**3.69**	**3.61**	**3.55**
17	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45	2.41	2.38
	8.40	**6.11**	**5.18**	**4.67**	**4.34**	**4.10**	**3.93**	**3.79**	**3.68**	**3.59**	**3.52**	**3.45**
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34
	8.28	**6.01**	**5.09**	**4.58**	**4.25**	**4.01**	**3.85**	**3.71**	**3.60**	**3.51**	**3.44**	**3.37**
19	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43	2.38	2.34	2.31
	8.18	**5.93**	**5.01**	**4.50**	**4.17**	**3.94**	**3.77**	**3.63**	**3.52**	**3.43**	**3.36**	**3.30**
20	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40	2.35	2.31	2.28
	8.10	**5.85**	**4.94**	**4.43**	**4.10**	**3.87**	**3.71**	**3.56**	**3.45**	**3.37**	**3.30**	**3.23**
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25
	8.02	**5.78**	**4.87**	**4.37**	**4.04**	**3.81**	**3.65**	**3.51**	**3.40**	**3.31**	**3.24**	**3.17**
22	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35	2.30	2.26	2.23
	7.94	**5.72**	**4.82**	**4.31**	**3.99**	**3.76**	**3.59**	**3.45**	**3.35**	**3.26**	**3.18**	**3.12**
23	4.28	3.42	3.03	2.80	2.64	2.53	2.45	2.38	2.32	2.28	2.24	2.20
	7.88	**5.66**	**4.76**	**4.26**	**3.94**	**3.71**	**3.54**	**3.41**	**3.30**	**3.21**	**3.14**	**3.07**
24	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30	2.26	2.22	2.18
	7.82	**5.61**	**4.72**	**4.22**	**3.90**	**3.67**	**3.50**	**3.36**	**3.25**	**3.17**	**3.09**	**3.03**
25	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24	2.20	2.16
	7.77	**5.57**	**4.68**	**4.18**	**3.86**	**3.63**	**3.46**	**3.32**	**3.21**	**3.13**	**3.05**	**2.99**
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15
	7.72	**5.53**	**4.64**	**4.14**	**3.82**	**3.59**	**3.42**	**3.29**	**3.17**	**3.09**	**3.02**	**2.96**

கட்டின்மை எண்ணிக்கை
(f_1 Degrees of Freedom (for greater mean square)

14	16	20	24	30	40	50	75	100	200	500	∞	f_2
245 **6,142**	246 **6,169**	248 **6,208**	249 **6,234**	250 **6,258**	251 **6,286**	252 **6,302**	253 **6,323**	253 **6,334**	254 **6,352**	254 **6,361**	254 **6,366**	1
19.42 **99.43**	19.43 **99.44**	19.44 **99.45**	19.45 **99.46**	19.46 **99.47**	19.47 **99.48**	19.47 **99.48**	19.48 **99.49**	19.49 **99.49**	19.49 **99.49**	19.50 **99.50**	19.50 **99.50**	2
8.71 **26.92**	8.69 **26.83**	8.66 **26.69**	8.64 **26.60**	8.62 **26.50**	8.60 **26.41**	8.58 **26.35**	8.57 **26.27**	8.56 **26.23**	8.54 **26.18**	8.54 **26.14**	8.53 **26.12**	3
5.87 **14.24**	5.84 **14.15**	5.80 **14.02**	5.77 **13.93**	5.74 **13.83**	5.71 **13.74**	5.70 **13.69**	5.68 **13.61**	5.66 **13.57**	5.65 **13.52**	5.64 **13.48**	5.63 **13.46**	4
4.64 **9.77**	4.60 **9.68**	4.56 **9.55**	4.53 **9.47**	4.50 **9.38**	4.46 **9.29**	4.44 **9.24**	4.42 **9.17**	4.40 **9.13**	4.38 **9.07**	4.37 **9.04**	4.36 **9.02**	5
3.96 **7.60**	3.92 **7.52**	3.87 **7.39**	3.84 **7.31**	3.81 **7.23**	3.77 **7.14**	3.75 **7.09**	3.72 **7.02**	3.71 **6.99**	3.69 **6.94**	3.68 **6.90**	3.67 **6.88**	6
3.52 **6.35**	3.49 **6.27**	3.44 **6.15**	3.41 **6.07**	3.38 **5.98**	3.34 **5.90**	3.32 **5.85**	3.29 **5.78**	3.28 **5.75**	3.25 **5.70**	3.24 **5.67**	3.23 **5.65**	7
3.23 **5.56**	3.20 **5.48**	3.15 **5.36**	3.12 **5.28**	3.08 **5.20**	3.05 **5.11**	3.03 **5.06**	3.00 **5.00**	2.98 **4.96**	2.96 **4.91**	2.94 **4.88**	2.93 **4.86**	8
3.02 **5.00**	2.98 **4.92**	2.93 **4.80**	2.90 **4.73**	2.86 **4.64**	2.82 **4.56**	2.80 **4.51**	2.77 **4.45**	2.76 **4.41**	2.73 **4.36**	2.72 **4.33**	2.71 **4.31**	9
2.86 **4.60**	2.82 **4.52**	2.77 **4.41**	2.74 **4.33**	2.70 **4.25**	2.67 **4.17**	2.64 **4.12**	2.61 **4.05**	2.59 **4.01**	2.56 **3.96**	2.55 **3.93**	2.54 **3.91**	10
2.74 **4.29**	2.70 **4.21**	2.65 **4.10**	2.61 **4.02**	2.57 **3.94**	2.53 **3.86**	2.50 **3.80**	2.47 **3.74**	2.45 **3.70**	2.42 **3.66**	2.41 **3.62**	2.40 **3.60**	11
2.64 **4.05**	2.60 **3.98**	2.54 **3.86**	2.50 **3.78**	2.46 **3.70**	2.42 **3.61**	2.40 **3.56**	2.36 **3.49**	2.35 **3.46**	2.32 **3.41**	2.31 **3.38**	2.30 **3.36**	12
2.55 **3.85**	2.51 **3.78**	2.46 **3.67**	2.42 **3.59**	2.38 **3.51**	2.34 **3.42**	2.32 **3.37**	2.28 **3.30**	2.26 **3.27**	2.24 **3.21**	2.22 **3.18**	2.21 **3.16**	13
2.48 **3.70**	2.44 **3.62**	2.39 **3.51**	2.35 **3.43**	2.31 **3.34**	2.27 **3.26**	2.24 **3.21**	2.21 **3.14**	2.19 **3.11**	2.16 **3.06**	2.14 **3.02**	2.13 **3.00**	14
2.43 **3.56**	2.39 **3.48**	2.33 **3.36**	2.29 **3.29**	2.25 **3.20**	2.21 **3.12**	2.18 **3.07**	2.15 **3.00**	2.12 **2.97**	2.10 **2.92**	2.08 **2.89**	2.07 **2.87**	15
2.37 **3.45**	2.33 **3.37**	2.28 **3.25**	2.24 **3.18**	2.20 **3.10**	2.16 **3.01**	2.13 **2.96**	2.09 **2.89**	2.07 **2.86**	2.04 **2.80**	2.02 **2.77**	2.01 **2.75**	16
2.33 **3.35**	2.29 **3.27**	2.23 **3.16**	2.19 **3.08**	2.15 **3.00**	2.11 **2.92**	2.08 **2.86**	2.04 **2.79**	2.02 **2.76**	1.99 **2.70**	1.97 **2.67**	1.96 **2.65**	17
2.29 **3.27**	2.25 **3.19**	2.19 **3.07**	2.15 **3.00**	2.11 **2.91**	2.07 **2.83**	2.04 **2.78**	2.00 **2.71**	1.98 **2.68**	1.95 **2.62**	1.93 **2.59**	1.92 **2.57**	18
2.26 **3.19**	2.21 **3.12**	2.15 **3.00**	2.11 **2.92**	2.07 **2.84**	2.02 **2.76**	2.00 **2.70**	1.96 **2.63**	1.94 **2.60**	1.91 **2.54**	1.90 **2.51**	1.88 **2.49**	19
2.23 **3.13**	2.18 **3.05**	2.12 **2.94**	2.08 **2.86**	2.04 **2.77**	1.99 **2.69**	1.96 **2.63**	1.92 **2.56**	1.90 **2.53**	1.87 **2.47**	1.85 **2.44**	1.84 **2.42**	20
2.20 **3.07**	2.15 **2.99**	2.09 **2.88**	2.05 **2.80**	2.00 **2.72**	1.96 **2.63**	1.93 **2.58**	1.89 **2.51**	1.87 **2.47**	1.84 **2.42**	1.82 **2.38**	1.81 **2.36**	21
2.18 **3.02**	2.13 **2.94**	2.07 **2.83**	2.03 **2.75**	1.98 **2.67**	1.93 **2.58**	1.91 **2.53**	1.87 **2.46**	1.84 **2.42**	1.81 **2.37**	1.80 **2.33**	1.78 **2.31**	22
2.14 **2.97**	2.10 **2.89**	2.04 **2.78**	2.00 **2.70**	1.96 **2.62**	1.91 **2.53**	1.88 **2.48**	1.84 **2.41**	1.82 **2.37**	1.79 **2.32**	1.77 **2.28**	1.76 **2.26**	23
2.13 **2.93**	2.09 **2.85**	2.02 **2.74**	1.98 **2.66**	1.94 **2.58**	1.89 **2.49**	1.86 **2.44**	1.82 **2.36**	1.80 **2.33**	1.76 **2.27**	1.74 **2.23**	1.73 **2.21**	24
2.11 **2.89**	2.06 **2.81**	2.00 **2.70**	1.96 **2.62**	1.92 **2.54**	1.87 **2.45**	1.84 **2.40**	1.80 **2.32**	1.77 **2.29**	1.74 **2.23**	1.72 **2.19**	1.71 **2.17**	25
2.10 **2.86**	2.05 **2.77**	1.99 **2.66**	1.95 **2.58**	1.90 **2.50**	1.85 **2.41**	1.82 **2.36**	1.78 **2.28**	1.76 **2.25**	1.72 **2.19**	1.70 **2.15**	1.69 **2.13**	26

கட்டின்மை எண்ணிக்கை

f_1 Degrees of Freedom (for greater mean square)

f_2	1	2	3	4	5	6	7	8	9	10	11	12
27	4.21 7.68	3.35 5.49	2.96 4.60	2.73 4.11	2.57 3.79	2.46 3.56	2.37 3.39	2.30 3.26	2.25 3.14	2.20 3.06	2.16 2.98	2.13 2.93
28	4.20 7.64	3.34 5.45	2.95 4.57	2.71 4.07	2.56 3.76	2.44 3.53	2.36 3.36	2.29 3.23	2.24 3.11	2.19 3.03	2.15 2.95	2.12 2.90
29	4.18 7.60	3.33 5.42	2.93 4.54	2.70 4.04	2.54 3.73	2.43 3.50	2.35 3.33	2.28 3.20	2.22 3.08	2.18 3.00	2.14 2.92	2.10 2.87
30	4.17 7.56	3.32 5.39	2.92 4.51	2.69 4.02	2.53 3.70	2.42 3.47	2.34 3.30	2.27 3.17	2.21 3.06	2.16 2.98	2.12 2.90	2.09 2.84
32	4.15 7.50	3.30 5.34	2.90 4.46	2.67 3.97	2.51 3.66	2.40 3.42	2.32 3.25	2.25 3.12	2.19 3.01	2.14 2.94	2.10 2.86	2.07 2.80
34	4.13 7.44	3.28 5.29	2.88 4.42	2.65 3.93	2.49 3.61	2.38 3.38	2.30 3.21	2.23 3.08	2.17 2.97	2.12 2.89	2.08 2.82	2.09 2.75
36	4.11 7.39	3.26 5.25	2.86 4.38	2.63 3.89	2.48 3.58	2.36 3.35	2.28 3.18	2.21 3.04	2.15 2.94	2.10 2.86	2.06 2.78	2.03 2.72
38	4.10 7.35	3.25 5.21	2.85 4.34	2.62 3.86	2.46 3.54	2.35 3.32	2.26 3.15	2.19 3.02	2.14 2.91	2.09 2.82	2.05 2.75	2.02 2.69
40	4.08 7.31	3.23 5.18	2.84 4.31	2.61 3.83	2.45 3.51	2.34 3.29	2.25 3.12	2.18 2.99	2.12 2.88	2.07 2.80	2.04 2.73	2.00 2.66
42	4.07 7.27	3.22 5.15	2.83 4.29	2.59 3.80	2.44 3.49	2.32 3.26	2.24 3.10	2.17 2.96	2.11 2.86	2.06 2.77	2.02 2.70	1.99 2.64
44	4.06 7.24	3.21 5.12	2.82 4.26	2.58 3.78	2.43 3.46	2.31 3.24	2.23 3.07	2.16 2.94	2.10 2.84	2.05 2.75	2.01 2.68	1.98 2.63
46	4.05 7.21	3.20 5.10	2.81 4.24	2.57 3.76	2.42 3.44	2.30 3.22	2.22 3.05	2.14 2.92	2.09 2.82	2.04 2.73	2.00 2.66	1.97 2.60
48	4.04 7.19	3.19 5.08	2.80 4.22	2.56 3.74	2.41 3.42	2.30 3.20	2.21 3.04	2.14 2.90	2.08 2.80	2.03 2.71	1.99 2.64	1.96 2.58
50	4.03 7.17	3.18 5.06	2.79 4.20	2.56 3.72	2.40 3.41	2.29 3.18	2.20 3.02	2.13 2.88	2.07 2.78	2.02 2.70	1.98 2.62	1.95 2.56
55	4.02 7.12	3.17 5.01	2.78 4.16	2.54 3.68	2.38 3.37	2.27 3.15	2.18 2.98	2.11 2.85	2.05 2.75	2.00 2.66	1.97 2.59	1.93 2.53
60	4.00 7.08	3.15 4.98	2.76 4.13	2.52 3.65	2.37 3.34	2.25 3.12	2.17 2.95	2.10 2.82	2.04 2.72	1.99 2.63	1.95 2.56	1.92 2.50
65	3.99 7.04	3.14 4.95	2.75 4.10	2.51 3.62	2.36 3.31	2.24 3.09	2.15 2.93	2.08 2.79	2.02 2.70	1.98 2.61	1.94 2.54	1.90 2.47
70	3.98 7.01	3.13 4.92	2.74 4.08	2.50 3.60	2.35 3.29	2.23 3.07	2.14 2.91	2.07 2.77	2.01 2.67	1.97 2.59	1.93 2.61	1.89 2.45
80	3.96 6.96	3.11 4.88	2.72 4.04	2.48 3.56	2.33 3.25	2.21 3.04	2.12 2.87	2.06 2.74	1.99 2.64	1.95 2.55	1.91 2.48	1.88 2.41
100	3.94 6.90	3.09 4.82	2.70 3.98	2.46 3.51	2.30 3.20	2.19 2.99	2.10 2.82	2.03 2.69	1.97 2.59	1.92 2.51	1.88 2.43	1.85 2.36
125	3.92 6.84	3.07 4.78	2.68 3.94	2.44 3.47	2.29 3.17	2.17 2.95	2.08 2.79	2.01 2.65	1.95 2.56	1.90 2.47	1.86 2.40	1.83 2.33
150	3.91 6.81	3.06 4.75	2.67 3.91	2.43 3.44	2.27 3.14	2.16 2.92	2.07 2.76	2.00 2.62	1.94 2.53	1.89 2.44	1.85 2.37	1.82 2.30
200	3.89 6.76	3.04 4.71	2.65 3.88	2.41 3.41	2.26 3.11	2.14 2.90	2.05 2.73	1.98 2.60	1.92 2.50	1.87 2.41	1.83 2.34	1.80 2.28
400	3.86 6.70	3.02 4.66	2.62 3.83	2.39 3.36	2.23 3.06	2.12 2.85	2.03 2.69	1.96 2.55	1.90 2.46	1.85 2.37	1.81 2.29	1.78 2.23
1,000	3.85 6.66	3.00 4.62	2.61 3.80	2.38 3.34	2.22 3.04	2.10 2.82	2.02 2.66	1.96 2.53	1.89 2.43	1.84 2.34	1.80 2.26	1.76 2.20
∞	3.84 6.64	2.99 4.60	2.60 3.78	2.37 3.32	2.21 3.02	2.09 2.80	2.01 2.64	1.94 2.51	1.88 2.41	1.83 2.32	1.79 2.24	1.75 2.18

கட்டின்மை எண்ணிக்கை
(f_1 Degrees of freedom for greater mean square)

14	16	20	24	30	40	50	75	100	200	500	∞	f_2
2.08	2.03	1.97	1.93	1.88	1.84	1.80	1.76	1.74	1.71	1.68	1.67	27
2.83	2.74	2.63	2.55	2.47	2.38	2.33	2.25	2.21	2.16	2.12	2.10	
2.06	2.02	1.96	1.91	1.87	1.81	1.78	1.75	1.72	1.69	1.67	1.65	28
2.80	2.71	2.60	2.52	2.44	2.35	2.30	2.22	2.18	2.13	2.09	2.06	
2.05	2.00	1.94	1.90	1.85	1.80	1.77	1.73	1.71	1.68	1.65	1.64	29
2.77	2.68	2.57	2.49	2.41	2.32	2.27	2.19	2.15	2.10	2.06	2.03	
2.04	1.99	1.93	1.89	1.84	1.79	1.76	1.72	1.69	1.66	1.64	1.62	30
2.74	2.66	2.55	2.47	2.38	2.29	2.24	2.16	2.13	2.07	2.03	2.01	
2.02	1.97	1.91	1.86	1.82	1.76	1.74	1.69	1.67	1.64	1.61	1.59	32
2.70	2.62	2.51	2.42	2.34	2.25	2.20	2.12	2.08	2.02	1.98	1.96	
2.00	1.95	1.89	1.84	1.80	1.74	1.71	1.67	1.64	1.61	1.59	1.57	34
2.66	2.58	2.47	2.38	2.30	2.21	2.15	2.08	2.04	1.98	1.94	1.91	
1.98	1.93	1.87	1.82	1.78	1.72	1.69	1.65	1.62	1.59	1.56	1.55	36
2.62	2.54	2.43	2.35	2.26	2.17	2.12	2.04	2.00	1.94	1.90	1.67	
1.96	1.92	1.85	1.80	1.76	1.71	1.67	1.63	1.60	1.57	1.54	1.53	38
2.59	2.51	2.40	2.32	2.22	2.14	2.08	2.00	1.97	1.90	1.86	1.84	
1.95	1.90	1.84	1.79	1.74	1.69	1.66	1.61	1.59	1.55	1.53	1.51	40
2.56	2.49	2.37	2.29	2.20	2.11	2.05	1.97	1.94	1.88	1.84	1.81	
1.94	1.89	1.82	1.78	1.73	1.68	1.64	1.60	1.57	1.54	1.51	1.49	42
2.54	2.46	2.35	2.26	2.17	2.08	2.02	1.94	1.91	1.85	1.80	1.78	
1.92	1.88	1.81	1.76	1.72	1.66	1.63	1.58	1.56	1.52	1.50	1.48	44
2.52	2.44	2.32	2.24	2.15	2.06	2.00	1.92	1.88	1.82	1.78	1.75	
1.91	1.87	1.80	1.75	1.71	1.65	1.62	1.57	1.54	1.51	1.48	1.46	46
2.50	2.42	2.30	2.22	2.13	2.04	1.98	1.90	1.86	1.80	1.76	1.72	
1.90	1.86	1.79	1.74	1.70	1.64	1.61	1.56	1.53	1.50	1.47	1.45	48
2.48	2.40	2.28	2.20	2.11	2.02	1.96	1.88	1.84	1.78	1.73	1.70	
1.90	1.85	1.78	1.74	1.69	1.63	1.60	1.55	1.52	1.48	1.46	1.44	50
2.46	2.39	2.26	2.18	2.10	2.00	1.94	1.86	1.82	1.76	1.71	1.68	
1.88	1.83	1.76	1.72	1.67	1.61	1.58	1.52	1.50	1.46	1.43	1.41	55
2.43	2.35	2.23	2.15	2.06	1.96	1.90	1.82	1.78	1.71	1.66	1.64	
1.86	1.81	1.75	1.70	1.65	1.59	1.56	1.50	1.48	1.44	1.41	1.39	60
2.40	2.32	2.20	2.12	2.03	1.93	1.87	1.79	1.74	1.68	1.63	1.60	
1.85	1.80	1.73	1.68	1.63	1.57	1.54	1.49	1.46	1.42	1.39	1.37	65
2.37	2.30	2.18	2.09	2.00	1.90	1.84	1.76	1.71	1.64	1.60	1.56	
1.84	1.79	1.72	1.67	1.62	1.56	1.53	1.47	1.45	1.40	1.37	1.35	70
2.35	2.28	2.15	2.07	1.98	1.88	1.82	1.74	1.69	1.62	1.56	1.53	
1.82	1.77	1.70	1.65	1.60	1.54	1.51	1.45	1.42	1.38	1.35	1.32	80
2.32	2.24	2.11	2.03	1.94	1.84	1.78	1.70	1.65	1.57	1.52	1.49	
1.79	1.75	1.68	1.63	1.57	1.51	1.48	1.42	1.39	1.34	1.30	1.28	100
2.26	2.19	2.06	1.98	1.89	1.79	1.73	1.64	1.59	1.51	1.46	1.43	
1.77	1.72	1.65	1.60	1.55	1.49	1.45	1.39	1.36	1.31	1.27	1.25	125
2.23	2.15	2.03	1.94	1.85	1.75	1.68	1.59	1.54	1.46	1.40	1.37	
1.76	1.71	1.64	1.59	1.54	1.47	1.44	1.37	1.34	1.29	1.25	1.22	150
2.20	2.12	2.00	1.91	1.83	1.72	1.66	1.56	1.51	1.43	1.37	1.33	
1.74	1.69	1.62	1.57	1.52	1.45	1.42	1.35	1.32	1.26	1.22	1.19	200
2.17	2.09	1.97	1.88	1.79	1.69	1.62	1.53	1.48	1.39	1.33	1.28	
1.72	1.67	1.60	1.54	1.49	1.42	1.38	1.32	1.28	1.22	1.16	1.13	400
2.12	2.04	1.92	1.84	1.74	1.64	1.57	1.47	1.42	1.32	1.24	1.19	
1.70	1.65	1.58	1.53	1.47	1.41	1.36	1.30	1.26	1.19	1.13	1.08	1,000
2.09	2.01	1.89	1.81	1.71	1.61	1.54	1.44	1.38	1.28	1.19	1.11	
1.69	1.64	1.57	1.52	1.46	1.40	1.35	1.28	1.24	1.17	1.11	1.00	∞
2.07	1.99	1.87	1.79	1.69	1.59	1.52	1.41	1.36	1.25	1.15	1.00	

This table is reprinted by permission from table number A 14 of Statistical Methods 6th edition by George W. Snedecor and William G. Cochran, © 1967 by The Iowa State University Press, Ames. Iowa, U. S. A.

TABLE—5

Random Numbers

Random Digits				
7442	4083	4593	5931	9844
5968	3047	0472	9141	4220
9478	9224	8471	0926	0310
9549	5236	2903	7241	8677
1763	1972	7035	0591	6761
8608	3298	1815	7279	0553
5707	5434	0921	9622	9730
9012	7487	1766	1747	1157
2797	0087	0151	4483	3227
5002	8449	8547	7759	8537
8053	8671	8718	2844	4808
8725	9122	8674	5661	1964
2947	6857	0393	6260	7916
9545	2086	4017	0290	8043
3173	3195	4195	4006	8901
7907	6404	5098	9805	9700
0660	1290	1481	0170	5976
8607	5404	9335	2698	9447
2702	9406	6788	7624	6850
4092	5306	0210	4018	9752
1962	3800	0947	1358	1499
8571	5495	3948	4556	8917
8583	4396	3916	7627	7870
6958	6450	6274	4092	7403
7813	3137	1042	4513	3793
6202	5081	7437	9792	7482
3602	3802	9585	0233	1125
7362	0540	9927	6753	7820
8594	8665	1428	7110	3650
8401	8717	6042	3832	0579

Random Digits

2315	3229	2885	5354	8719
9038	5181	1297	3073	3897
9108	1907	0961	5651	8249
6973	3335	0603	4089	5798
9105	8156	8160	1915	0154
0059	6531	4733	3270	3278
9296	7099	4717	7110	3786
9468	6225	9161	1447	6750
0788	5580	8934	6084	0462
5997	0660	3514	0122	7510
3540	0545	6249	2134	8217
7917	5174	8048	1128	5968
5078	2220	8988	8596	9655
6378	4422	4235	8116	3074
0979	6571	3607	0119	2188
4918	7024	3667	0480	1029
6817	5761	3709	4728	5168
4620	1539	0915	6348	2908
9444	8857	2542	3169	5838
0865	2948	0117	9410	3168
2923	1870	4410	1107	6502
7564	4456	9381	5450	4201
9243	9996	5474	4545	5976
4698	6851	7388	6221	7690
1679	7502	1501	8182	2791
2452	4566	0063	2998	6206
7718	1595	1663	9712	1949
9581	8551	2222	8934	6427
2606	0638	5831	8923	3804
6225	3976	9414	6561	1843
5525	2406	3403	4926	7858
0435	1737	5504	0171	9765
7402	4174	6749	1115	1256
7578	3127	9779	1225	3014
8080	6396	5680	9512	4864

Random Digits

2862	2536	6924	2657	4592
0263	7566	0818	1617	5577
7825	2207	8947	0037	6266
4153	6787	3179	5191	8773
0169	7207	3942	7479	7106
3477	4956	5529	2512	5832
7520	9583	1316	0525	4829
0303	4653	1699	9939	7449
7735	2502	2005	0631	9888
8153	2683	5772	4839	9714
0165	8156	1172	2637	4896
7849	5240	8311	9400	4421
6140	2003	7009	2407	8875
6546	9212	9380	8736	6120
2471	7972	4706	3986	9176
8094	7747	6006	2536	8856
3745	6766	7221	9560	7036
0835	3641	1638	3464	1767
1601	1143	7272	3988	8356
7797	4121	9603	5723	4630
7620	8310	9500	7116	6259
2096	5270	0793	3950	2722
6803	7016	1055	6396	7754
3566	9310	2604	8607	4765
6428	0086	6245	3247	5707
3755	5044	0128	4083	9181
6487	2778	8138	6776	0208
4163	1822	5656	9019	0138
8330	4855	5576	5727	0232
1879	0353	9354	2673	8744
0559	8989	0700	8672	9191
2833	3250	5106	6[illegible]05	9458
8351	1668	6054	4681	6656
6780	8083	0621	1482	1933
0759	3015	8887	3523	1628

Random Digits

1931	9402	5236	2853	9537
1377	3459	1811	9418	1109
2455	2433	8860	4592	1811
9188	7751	2595	6431	4409
1706	9388	9554	1170	3857
5160	3553	4946	9352	5504
5264	2153	0571	0809	0978
4854	8776	2078	7518	3425
8205	9414	9621	8669	1445
5669	6598	3196	1572	0455
8171	0291	0603	4675	0779
4520	4584	5714	8122	5029
7664	6247	8362	1257	5269
9477	5870	6425	1725	9792
5549	0593	5761	7200	7229
7619	3749	9121	2185	9335
0925	5792	1040	5806	9636
3591	6713	5325	7485	2406
2237	1222	3947	1228	2708
7847	6217	0857	8229	5609
8633	2617	9176	9602	4807
6632	8056	1091	7158	7303
2612	7936	1453	4812	1742
9436	1681	0851	3488	8815
0418	2487	5583	9032	6507
6853	3757	0171	5943	1145
8347	7044	4640	6832	2445
5182	6263	1224	9863	6751
9215	3992	4874	8002	5959
5588	3456	9602	5260	6578
3996	0415	7015	9210	0974
7346	9400	3292	8165	3206
8621	4185	6727	2770	1227
9399	5575	1562	5821	9824
4334	0347	4893	2025	5590

Random Digits

8091	0536	7522	5409	1463
9627	3311	2010	2525	3142
0086	3501	4916	2511	1274
3753	0174	7934	3483	9210
2740	32[illegible]9	3054	9991	3778
3919	6871	5685	8147	7310
4577	7897	2757	5992	7398
0215	7254	5378	3861	3448
5807	7948	1774	6836	1786
1910	9653	1214	3921	5298
3990	1310	9338	2601	5571
5967	8941	7987	3335	7579
5872	1143	9183	6911	2247
7240	1827	3281	0705	4479
2268	4227	5844	0700	6907
7269	6131	8780	3417	7278
4084	9096	4047	6775	0876
7128	3636	6561	7522	0359
5301	5403	5456	0501	4511
8554	0346	6251	3577	4146
3434	0188	5665	7779	7179
4872	7870	2335	2874	9393
0084	8827	9479	8342	0053
2861	4574	5813	5903	7161
8618	0340	3381	7579	6359
0319	2699	8036	1090	3805
7035	5277	7340	8515	4225
8696	6496	4889	2897	3316
4909	0348	8735	3604	9959
8126	8571	2532	9355	7563
0138	0384	6711	2384	0072
9700	2196	4076	3710	3372
1775	8324	9646	0611	1048
9163	4714	7888	3577	6596
3195	1040	2022	3193	9196

Random Digits

2080	4196	3575	5700	7967
7687	8415	1595	9636	4605
9494	5221	1325	7317	1022
2392	2820	8533	0629	3771
8334	2352	7113	2291	7312
1424	7850	4531	0133	5519
9735	3042	8409	7053	5364
1559	4838	7198	9249	1395
5598	9985	8170	3367	6928
9668	6670	0097	0686	6311

This table is extracted from table 19.1 of C. R. Rao, S. K. Mitra and A. Matthai 'Formulae and tables for Statistical work'—Statistical Publishing Society—Calcutta—with the permission of the Authors and Publishers.

LIST OF REFERENCES

1. Anderson, R. L. and Bancroft, T. A., *Statistical Theory in Research*, Newyork, Mc Graw-Hill, 1952.

2. Arthur, E. Mace, *Sample-size determination*, Reirihold publishing society, Newyork, 1964.

3. Bailey, Norman T. J. *Statistical Methods in Biology*, the English University Press, London, 1969.

4. Bishop, O. N. *Statistics for Biology*, Longmans, 1966.

5. Dixon, W. J. and Massey, F. J. Jr. *Introduction to Statistical Analysis*, Newyork, Mc Graw-Hill, 1942.

6. Fisher, Sir Ronald (R. A.) *The Design of experiments* 4th ed., Edinburgh and London, Oliver and Boyd Ltd., 1942.

7. Fisher, Sir Ronald (R.A.,) *Statistical Methods for Research Workers*, 10th Ed., Edinburgh and London, Oliver and Boyd., Ltd., 1948.

8. Fisher, Sri Ronald (R.A.) and Yates, F., *Statistical tables for Biological, Agricultural and Medical Research*, 6th Ed., Edinburgh and London, Oliver and Boyd, Ltd., 1963.

9. Frederick C. Mills, *Statistical Methods*, Henry Holt and Company, Newyork, 1955.

10. Frederick E. Croxton and Cowdon D. J. *Applied General Statistics*, Sir Isace Pitmans Sons Ltd., London, 1956.

11. Gouldon, C. H. *Methods of Statistical Analysis*, 2nd Ed., Wiley, Newyork, 1952.

12. Gunnar Dahlbage, *Statistical Methods for Medical and Biological Students*, George Ollen and Unwin Ltd., 1940.

13. Hod. P. G., *Introduction to Mathematical Statistics*, 2nd Ed., Newyork, Wiley, 1954.

14. John G. Peatman, *Introduction to Applied Statistics*, Harper and Row, Newyork, 1964.

15. Kendall, M. G. and Buckland, W. R., *A Dictionary of Statistical Terms*, Oliver and Boyd, London, 1957.

16. Mood, A. M., *Introduction to the Theory of Statistics*, Newyork, Mc Graw Hill, 1950.

17. Oscar Kempthorne, *An Introduction to Venetic Statistics*, John Wiley, Newyork, 1957.

18. Oscar Kempthorne and Others, edited by *Statistics and Mathematics in Biology*, The Iowa State College Press, Iowa, 1954.

19. Pearson, E. S. and Hontley, H. O., *Biometrika Table for Statisticians*, Vol. 1, Cambridge University Press, 1954.

20. Radhakrishna Rao. e, *Advanced Statistical Methods in Biometric Research*, John Wiley and Sons, *Inc.*, Newyork, 1952.

21. Radhakrishna Rao. e, Mitra, S. K., and Mathari, A., *Formulae and Table for Statistical Work*, Statistical Publishing Society, Calcutta, 1966.

22. Rosander, A. L., *Elementary Principles of Statistics*, D Van Nostrand *Inc*, Newyork, 1957.

23. Raymond Pearl, *Introduction to Biometry and Statistics*, W. B. Saunders Company, 1930.

24. Snedeer, George. E., and Cochran, William G., *Statistical Methods*, 6th Ed., The Iowa State University Press, Ames, Iowa., 1967.

25. Tolbot H. Waterman, *Theoritical and Mathemetical Biology*, Blaisdell Publishing Company, Newyork, 1965.

26. Weatherburn, C. E., *A First Course in Mathematical Statistics*, The English Language book Society, 1961.

27. Yule, G. U. and Kendal, M. G., *An Introduction to the Theory of Statistics*, 14th Ed., Newyork, Hafner, 1950.

28. சண்முகசுந்தரம், கோ., புள்ளியியல் முறைகள், பாகம் 1 தமிழ் வெளியீட்டுக் கழகம், தமிழ்நாடு அரசாங்கம் 1964.

29. ராஜகோபாலன், K. R., புள்ளியியல் முறைகள், பாகம் 2 தமிழ் வெளியீட்டுக் கழகம், தமிழ்நாடு அரசாங்கம், 1966.

30. வைத்தியநாதன், சு., புள்ளியியல்—அறிமுகம், கல்லூரி நூல் வெளியீட்டு இயக்குநரகம், தமிழ்நாட்டுப் பாடநூல் நிறுவனம், மறுபதிப்பு 1970.

கலைச்சொற்கள்

(ஆங்கிலம்—தமிழ்)

A

Aggregate	—	திரள், மொத்தம்
Analysis	—	ஆய்வு, பகுப்பாய்வு
Analysis of variance	—	மாறுபாட்டளவை ஆய்வு
Angle	—	கோணம்
Applied Statistics	—	பயன்முறைப் புள்ளியியல்
Arithmetic mean	—	கூட்டுச் சராசரி
Arrays	—	வரிசை
Ascending	—	ஏறுகின்ற
Ascending order	—	ஏறுவரிசை
Association	—	இயைபு
Association coefficient of	—	இயைபு கெழு
,, illusory	—	போலி இயைபு
,, negative	—	எதிரிடை இயைபு
,, of attributes	—	பண்புகளின் இயைபு
,, partial	—	பகுதி இயைபு
Attribute	—	பண்பு
,, negative	—	எதிரிடைப் பண்பு (பண் பின்மை)
,, positive	—	நேரிடைப் பண்பு (பண் பமையப் பெற்ற நிலை)
Average	—	சராசரி (பொதுப்பெயர்)
Axis	—	அச்சு
Asymmetry	—	சமச்சீரின்மை
Asymmetrical distribution	—	சமச்சீரில்லாப் பரவல்

B

Bell shaped	—	மணி வடிவமான
Best fit	—	மிகப் பொருத்தமான
Binomial	—	ஈருறுப்பு
,, co-efficient	—	ஈருறுப்புக் குணகம், ஈருறுப்புக் கெழு

Binomial distribution — ஈருறுப்புப் பரவல்
,, expansion — ஈருறுப்புத் தேற்றவிரிவு
,, expression — ஈருறுப்புக் கோவை
,, theorem — ஈருறுப்புத் தேற்றம்
Biometry — உயிர்ப் புள்ளியியல்
Birth rate — பிறப்பு வீதம்
Bivariate — இருமாறி

C

Calculate — கணக்கிடு
Calculation — கணிப்பு
Casual relationship — தற்செயல் தொடர்பு
Cause and effect relationship — காரண விளைவுத் தொடர்பு
Census — மக்கட்கணிப்பு
,, population — மக்கள் மதிப்பீடு
Cent — நூறு, சதம்
Central tendency — மையநிலைப் போக்கு
,, ,, measures of — மையநிலைப் போக்கு அளவைகள்
Child mortality rate — குழந்தை இறப்பு வீதம்
Chi-square — கை-வர்க்கம்
,, distribution — கை-வர்க்கப் பரவல்
,, table — கை-வர்க்கப் பட்டியல்
,, test — கை-வர்க்கச் சோதனை
Class — பிரிவு
,, frequency — பிரிவு நிகழ்வெண்
,, interval — பிரிவு இடைவெளி, இடைவெளிப் பிரிவு
,, interval exclusive — நுழைவு தவிர்க்கின்ற இடைவெளிப் பிரிவு
,, interval inclusive — உள்ளடக்கிய இடைவெளிப் பிரிவு
,, interval true — உண்மை இடைவெளிப் பிரிவு
,, interval uniform — சமஅளவு இடைவெளிப் பிரிவு
,, interval ununiform — சமஅளவில்லா இடைவெளிப் பிரிவு
,, limit — பிரிவு எல்லை
,, limit lower — பிரிவுக் கீழ்எல்லை
,, limit upper — பிரிவு மேல் எல்லை
,, mark — பிரிவெண் (*mid x*)

Classification		— பிரித்தமைத்தல்
,,	one way	— ஒருவழிப் பாகுபாடு
,,	two way	— இருவழிப் பாகுபாடு
,,	three way	— மூவழிப் பாகுபாடு
,,	manifold	— பல்வழிப் பாகுபாடு
Coefficient of association		— இயைபு கெழு
,,	of contingency	— இணைப்புக் கெழு
,,	of correlation	— உடன்தொடர்பு கெழு
,,	of dispersion	— சிதறல் கெழு
,,	of rank correlation	— மதிப்பிடத் தொடர்பு கெழு
,,	of regression	— மாறிகளின் தொடர்பு கெழு
,,	of skewness	— கோட்டக் கெழு
,,	of variation	— மாறுபாட்டுக் கெழு
Column		— நிரல், நிமிர்நிலை வரிசை
Combination		— சேர்மானம்
Confidence limits		— நம்பக எல்லை
,,	interval	— நம்பக இடைவெளி
Consistent		— இசைவு, பொருத்தமுடைமை
Constant		— மாறிலி
Constraint		— கட்டுமானம், தளை
Contingency		— இணைவு
,,	coefficient of	— இணைப்புக் கெழு
,,	table	— இணைப்புப் பட்டியல்
Continuous		— தொடர்ச்சியான
,,	variable	— தொடர்மாறி
Correction factor		— திருத்தக் காரணி
Correlation		— உடன்தொடர்பு, ஒட்டுறவு
,,	coefficient	— உடன்தொடர்பு கெழு
,,	multiple	— பல்தரத் தொடர்பு
,,	negative	— எதிர்த் தொடர்பு
,,	nonsense	— பொருளில்லாத் தொடர்பு
,,	partial	— பகுதித் தொடர்பு
,,	positive	— நேர் தொடர்பு
,,	rank	— மதிப்பிடத் தொடர்பு
,,	ratio	— உடன்தொடர்பு விகிதம்
,,	spurious	— போலித் தொடர்பு
,,	table	— உடன்தொடர்பு பட்டியல்
Covariation		— உடன்மாறுபாடு
Critical region		— மாறுநிலை மண்டிலம்

Cumulative	—	வளர், குவிவு
,, frequency	—	வளர்நிகழ்வெண்
,, ,, less than	—	கீழின வளர்நிகழ்வெண்
,, greater than	—	மேலின வளர்நிகழ்வெண்
Curve	—	வளைகோடு
,, frequency	—	நிகழ்வெண் வளைகோடு
,, normal	—	இயல்நிலை வளைகோடு
,, skewed	—	கோட்டமுடைட வளைகோடு
,, symmetrical	—	சமச்சீரான வளைகோடு

D

Data	—	விவரம், தொகுத்த செய்திகள்
Death rate	—	இறப்பு வீதம்
Degree	—	பாகை (கோணம்)
Degrees of freedom	—	கட்டின்மை எண்ணிக்கை
Dependent	—	சார்புடைய, சார்ந்த
,, variable	—	சார்புடை மாறி
Descending order	—	இறங்குவரிசை
Design	—	திட்டஅமைப்பு
,, latine square	—	லேடின் சதுரத் திட்ட அமைப்பு
,, randomised	—	சமவாய்ப்புத் திட்டஅமைப்பு
,, randomised block	—	சமவாய்ப்புக் கட்டுதிட்ட அமைப்பு
,, of experiments	—	சோதனைத் திட்ட அமைப்பு, செய்முறைத் திட்டம்
Deviation	—	விலக்கம்
,, standard	—	திட்ட விலக்கம்
,, quartile	—	கால்மான விலக்கம்
Die	—	பகடை
Discrete variable	—	தொடர்ச்சியற்ற மாறி, தனித்த மாறி
Dispersion	—	சிதறல்
,, measures of	—	சிதறல் அளவை
,, co-efficient	—	சிதறல் கெழு
Distribution	—	பரவல்
,, binomial	—	ஈருறுப்புப் பரவல்

Distribution chi-square (χ^2) — கை-வர்க்கப் பரவல்

,, continuous — தொடர்ச்சியற்ற பரவல்

,, discrete — தொடர்ச்சியற்ற பரவல்

,, 'F' — F-பரவல்

,, 'J'-shaped — 'J' வடிவப் பரவல்

,, normal — இயல்நிலைப் பரவல்

,, Poisson — பாய்ஸான் பரவல்

,, 'Student 't'' — t-பரவல்

,, skew — கோட்டமுடைப் பரவல், சமச்சீரற்ற பரவல்

,, symmetrical — சீரான பரவல்

,, 'U' shaped — 'U' வடிவப் பரவல்

E

Equation — சமன்பாடு

Error — பிழை

,, standard — திட்டப் பிழை

Estimate — மதிப்பிடு

Estimation — மதிப்பீடு

,, biased — பிறழ்ச்சியான மதிப்பீடு

,, consistent — முரண்படாத மதிப்பீடு

,, efficient — திறமுடைய மதிப்பீடு

,, interval — இடைவெளி மதிப்பீடு

,, point — புள்ளி மதிப்பீடு

,, sufficient — போதுமான மதிப்பீடு

,, unbiased — பிறழ்ச்சியற்ற மதிப்பீடு

Event — நிகழ்ச்சி

,, dependent — சார்புடை நிகழ்ச்சி

,, equally likely — சமவாய்ப்புடை நிகழ்ச்சி

,, exhaustive — பூரணத்தன்மைத்தான நிகழ்ச்சி

,, favourable — சாதக நிகழ்ச்சி

,, independent — சார்பற்ற நிகழ்ச்சி

,, mutually exclusive — ஒன்றையொன்று விலக்கும் தன்மைத்தான நிகழ்ச்சி

F

Factorial — காரணீயப் பெருக்கம்

Favourable event — சாதக நிகழ்ச்சி

Finite	—	முடிவுள்ள, வரையுற்ற
,, population	—	வரையுற்ற முழுமைத் தொகுதி
Fluctuation	—	ஏற்ற இறக்கம்
,, random	—	இயைபிலா ஏற்ற இறக்கம்
,, sampling	—	மாதிரி எடுத்தலின் ஏற்ற இறக்கம்
Formula	—	வாய்பாடு
Fraction	—	பின்னம்
Frequency	—	நிகழ்வெண்
,, cumulative	—	வளர்நிகழ்வெண்
,, cumulative greater than	—	மேலின வளர்நிகழ்வெண்
,, cumulative less than	—	கீழின வளர்நிகழ்வெண்
,, curve	—	நிகழ்வெண் வளைகோடு
,, distribution	—	நிகழ்வெண் பரவல்
,, observed	—	கண்டறிந்த நிகழ்வெண்
,, polygon	—	நிகழ்வெண் பலகோணம்
,, percentage	—	சதவீத நிகழ்வெண்
,, ratio	—	நிகழ்வெண் வீதம்
,, relative	—	ஒப்பீட்டு நிகழ்வெண்
,, table	—	நிகழ்வெண் பட்டியல்
,, theoretical	—	கருதுகோள் நிகழ்வெண்
,, table bivariable	—	இரட்டைமாறி நிகழ்வெண் பரவல்
,, function	—	நிகழ்வெண் சார்பு

G

Gap	—	பிளவு
Generalise	—	பொதுவிதி காண்
General law	—	பொது விதி
Geometric mean	—	பெருக்கற் சராசரி
Goodness of fit	—	பொருத்தச் செம்மை
Graph	—	வரைபடம்
Graphic presentation	—	வரைபட அமைப்பு
Greater than ogive	—	மேலின ஓகைவ், மேலின வளர் நிகழ்வெண் பரவல் வளைகோடு

H

Heterogeneous	— பல்படித்தான
Histogram	— நிகழ்வெண் செவ்வகப் படம்
Homogeneous	— ஒருபடித்தான
Hypothesis	— கருதுகோள்
,, null	— மறுக்கத்தகு கருதுகோள்
,, testing of	— கருதுகோள் சோதனை

I

Inconsistent	— இசைவற்ற
Infant mortality rate	— குழந்தை இறப்பு விகிதம்
Infinity	— எண்ணிலி, முடிவற்ற
Inteligence quotient	— அறிவுக் குறியீட்டெண்
Interval	— இடைவெளி
,, class	— இடைவெளிப் பிரிவு
,, confidence	— நம்பக இடைவெளி
Inter quartile range	— இடைக் கால்மான வீச்சு
,, semi	— அரை இடைக் கால்மான வீச்சு

L

Large sample	— பெரு அளவு மாதிரி
Less than ogive	— கீழின ஒகைவ், கீழின வளர் நிகழ்வெண் வளைகோடு
Life table	— வாழ்க்கைப் பட்டியல்
Limit	— எல்லை, வரம்பு
,, confidence	— நம்பக வரம்பு
,, lower	— கீழ் எல்லை
,, upper	— மேல் எல்லை
Linear	— நேர்கோட்டுக்குரிய, ஒரு படிக்குரிய
,, constraint	— முதலடுக்குக் கட்டுமானம், முதலடுக்குத் தளை
Lower limit	— கீழ் எல்லை
,, quartile	— கீழ்க் கால்மானம்

M

Magnitude	— பருமன், அளவு
,, ascending order of	— ஏறுவரிசை
,, descending order of	— இறங்குவரிசை

Mathematics — கணிதம்
,, advanced — உயர்கணிதம்
Mathematical expectation — கணக்கியல் எதிர்பார்ப்பு
Maximum — பெருமம்
Mean — கூட்டுச் சராசரி
Measure — அளவு
Median — இடைநிலை
Middle point — மையப்புள்ளி
Minimum — சிறுமம்
Mode — முகடு
Multiple correlation — பல்தரத் தொடர்பு

N

Normal curve of error — இயல்நிலைப் பிழை வளைகோடு
,, distribution — இயல்நிலைப் பரவல்
,, probability curve — இயல்நிலை நிகழ்வெண் வளைகோடு
Null hypothesis — மறுக்கத்தகு கருதுகோள்
Number — எண்
,, even — இரட்டைப்படை எண்
,, natural — இயற்கை யெண்
,, negative — குறை யெண்
,, odd — ஒற்றைப்படை யெண்
,, positive — கூட்டெண்

O

Observation — கண்டறிதல்
Ogive — ஓகைவ், வளர்நிகழ்வெண் வளைகோடு
,, greater than — மேலின ஓகைவ், மேலின நிகழ்வெண் வளைகோடு
,, less than — கீழின ஓகைவ், கீழின வளர் நிகழ்வெண் வளைகோடு
One tail test — ஒருவால்புறச் சோதனை
One way clasification — ஒருவழிப் பாகுபாடு

P

Parameter — முழுமைத் தொகுதி அளவை
Partial — பகுதி

Partial association	—	பகுதி இயைபு
,, correlation	—	பகுதித் தொடர்பு
Percent	—	சதவீதம்
Permutation	—	வரிசைமாற்றம்
Poisson distribution	—	பாய்ஸான் பரவல்
Population	—	முழுமைத் தொகுதி
,, artificial	—	செயற்கை உறுப்புகளைக் கொண்ட முழுமைத்தொகுதி
,, dynamic	—	இயங்கும் தன்மைத்தான முழுமைத்தொகுதி
,, finite	—	வரம்புடை முழுமைத்தொகுதி
,, hypothetical	—	கருதுகோள் முழுமைத்தொகுதி
,, infinite	—	எண்ணற்ற இனத்தொகுதி
,, known	—	சார்பு தெரிந்த முழுமைத் தொகுதி
,, real	—	உண்மையான உருப்பெற்ற முழுமைத் தொகுதி
,, static	—	நிலையான முழுமைத்தொகுதி
,, unknown	—	சார்பு தெரியாத முழுமைத் தொகுதி
Positive association	—	நேர் இயையு
,, number	—	நேர் எண், கூட்டெண்
,, skewness	—	நேர்கோட்டம்
Probability	—	நிகழ்திறன், யூக அளவை
,, conditional	—	நிபந்தனை நிகழ்திறன்
,, mathematical	—	கணக்கியல் நிகழ்திறன்
,, statistical	—	புள்ளியியல் நிகழ்திறன்
,, density function	—	நிகழ்திற அடர்த்திச் சார்பு
Probable error	—	நிகழ்திறப் பிழை, ஒரு பரவலின் நண்பாதி இடைவெளி
Proportion	—	விகித சமம்

Q

Qualitative data	—	பண்பின் அடிப்படையான விவரங்கள்
Quantitative data	—	அளவின் அடிப்படையான விவரங்கள்
Quartile	—	கால்மானம்
,, deviation	—	கால்மான விலக்கம்

Quartile first — முதல் கால்மானம்
,, lower — கீழ்க் கால்மானம்
,, third — மூன்றுவது கால்மானம்
,, upper — மேல் கால்மானம்
,, range — கால்மான வீச்சு
,, semi — அரைக் கால்மான வீச்சு
Quotient — ஈவு

R

Random — ராண்டம், இயைபிலா
,, fluctuation — ராண்டம் ஏற்ற இறக்கம்
,, numbers — ராண்டம் எண்கள்
,, sample — ராண்டம் மாதிரி, இயைபிலா மாதிரி
,, variable — ராண்டம் மாறி
Range — வீச்சு இடைவெளி
Rank — தரம், மதிப்பிடம்
,, correlation — மதிப்பிடத் தொடர்பு
,, ,, co-efficient of — மதிப்பிடத் தொடர்பு கெழு
Ratio — விகிதம்
Raw data — சீர்படா விவரங்கள்
Regression — மாறிகளின் தொடர்பு
,, co-efficient — மாறிகளின் தொடர்பு கெழு
,, equation — தொடர்பு சமன்பாடு
,, lines — மாறிகளின் தொடர்பு கோடு
,, line of x on y — x மாறியை y மாறியில் குறிக்கும் மாறிகளின் தொடர்பு கோடு
,, line of y on x — y மாறியை x மாறியில் குறிக்கும் மாறிகளின் தொடர்பு கோடு
Rejection region — மறுப்பு மண்டிலம்
Replical — திரும்பச் செய்தல்
Row — நிமிர்நிலை வரிசை

S

Sample — மாதிரி, கூறு
,, large — பெரு அளவு மாதிரி
,, simple random — இயைபிலா மாதிரி
,, size — மாதிரி அளவு

Sample small	— சிறிய அளவு மாதிரி
,, stratified random	— பகுப்புமுறை இயைபிலா மாதிரி
Sampling	— கூறெடுத்நல், மாநிரி அமைத்தல்
,, distribution	— மாதிரி அளவையின் (statistics) பரவல், கூறு பரவல்
,, fluctuation	— மாதிரி ஏற்ற இறக்கம்
,, method	— கூறெடுத்தல் முறைகள்
,, random	— இயைபிலாக் கூறெடுத்தல், சம வாய்ப்புக் கொடுத்தெடுக்கப் பட்ட மாதிரி
Scale	— அளவு திட்டம்
Scatter	— சிதறல்
,, diagram	— சிதறல் விளக்கப்படம்
Semi-interquartile range	— அரை இடைக் கால்மான வீச்சு
Sensitiveness	— கூருணர்வுத் திறம்
Sign	— குறி
Significance level	— சிறப்பு வரம்பு
,, test	— சிறப்புக் காணும் சோதனை
Skew	— கோட்டமுடைய, சமச்சீரற்ற
Skewe1 distribution	— கோட்டமுடைப் பரவல்
Skewness co-efficient of	— கோட்டக் கெழு
Skewness measures of	— கோட்ட அளவை
Slope	— சரிவு
Standard deviation	— திட்ட விலக்கம்
Standard error	— திட்டப் பிழை
Standard unit	— திட்ட அலகு
Statistic	— புள்ளியியல் அளவை, மாதிரி அளவை
Statistics (as data)	— புள்ளி விவரங்கள்
Statistics (as science)	— புள்ளியியல்
Statistical inference	— புள்ளியியல் அனுமானம்
Symmetrical	— சமச்சீரான
curve	— சீரான வளைகோடு
distribution	— சீரான பரவல்

T

Table	— பட்டியல்
contingency	— இணைப்புப் பட்டியல்

Table correlation — உடன்தொடர்பு பட்டியல்
,, life — வாழ்க்கைப் பட்டியல்
Tally mark — சரிபார்க்கும் குறி
Test — சோதனை
,, goodness of fit — இனச் சிறப்புநிலைச் சோதனை
,, of homogeneity — ஒருபடித்தான சோதனை
,, of independence — தொடர்பற்றநிலைச் சோதனை
,, staistic — சோதனை அளவை
True class interval — உண்மை இடைவெளிப் பிரிவு

U

Unbiased — பிறழ்ச்சியற்ற
Unit — அலகு
Universe — முழுமை
Unknown — மதிப்புத் தெரியாத, சார்பு தெரியாத
Unknown population — சார்பு தெரியாத முழுமைத் தொகுதி
Upper limit — மேல் எல்லை
,, quartile — மேல் கால்மானம்

V

Variable — மாறி
,, continuous — தொடர்மாறி
,, dependent — சார்புடை மாறி
,, discrete — தொடர்ச்சியற்ற மாறி
,, independent — சார்பற்ற மாறி
Variance — மாறுபாட்டளவை
Variance-analysis of — மாறுபாட்டளவை ஆய்வு
Variation — மாறுபாடு
,, measures of — மாறுபாட்டளவைகள்
,, co-efficient of — மாறுபாட்டுக் கெழு

X

X-axis — *X*-அச்சு

Y

Y-axis — *Y*-அச்சு

கலைச்சொற்கள்

(தமிழ்—ஆங்கிலம்)

அ

அச்சு — Axis
அலகு — Unit
அளவின் அடிப்படையான விவரங்கள் — Quantitative data
அளவு — Magnitude
அளவு திட்டம் — Scale
அளவை — Measure
அறிவுக் குறியீட்டெண் — Intelligent quotient

ஆ

ஆய்வு — Analysis

இ

இசைவற்ற — Inconsistent
இசைவு — Consistent
இடைக் கால்மான வீச்சு — Interquartile range
இடைநிலை — Median
இடைவெளி; — Interval
இடைவெளிப் பிரிவு — Class interval
இடைவெளி மதிப்பீடு — Interval estimation
இணைச் சிறப்புநிலைச் சோதனை — Goodness of fit
இணைப்புக் கெழு — Co-efficient of contingency
இணைப்புப் பட்டியல் — Contingency table
இணைவு — Contingency
இயல்நிலைப் பிழை வளைகோடு — Normal curve of error
இயல்நிலை நிகழ்திற வளைகோடு — Normal probability curve
இயங்கும் தன்மைத்தான முழுமைத் தொகுதி — Dynamic population
இயற்கை எண் — Natural number

இயைபு	—	Association
இயைபு கெழு	—	Co-efficient of association
இயைபிலா ஏற்ற இறக்கம்	—	Random fluctuation
இயைபிலாக் கூறெடுத்தல்	—	Random sample
இயைபிலா மாறி	—	Random variable
இரட்டைப்படை எண்	—	Even number
இரட்டை மாறி, இருமாறி	—	Bivariate
இரட்டை மாறி நிகழ்வெண் பரவல்	—	Bivariate frequency table
இருவழிப் பாகுபாடு	—	Two way classification
இறங்குவரிசை	—	Descending order
இறப்பு விகிதம்	—	Death rate

ஈ

ஈருறுப்பு	—	Binomial
ஈருறுப்புக் குணகம்	—	Binomial co-efficient
ஈருறுப்புக் கெழு	—	Binomial co-efficient
ஈருறுப்புக் கோவை	—	Binomial expression
ஈருறுப்புத் தேற்றம்	—	Binomial theorem
ஈருறுப்புத் தேற்ற விரிவு	—	Binomial expansion
ஈருறுப்புப் பரவல்	—	Binomial distribution
ஈவு	—	Quotient

உ

உடன்தொடர்பு	—	Correlation
உடன்தொடர்பு கெழு	—	Correlation co-efficient
உடன்தொடர்பு பட்டியல்	—	Correlation table
உடன்தொடர்பு விகிதம்	—	Correlation ratio
உடன்மாறுபாடு	—	Covariation
உண்மை இடைவெளிப் பிரிவு	—	True class interval
உண்மை உருப்பெற்ற முழுமைத் தொகுதி	—	Real population
உயர்கணிதம்	—	Advanced mathematics
உயிர்ப் புள்ளியியல்	—	Biometry
உள்ளடக்கிய இடைவெளிப் பிரிவு	—	Inclusive class interval

எ

எண் — Number
எண்ணிலி — Infinity
எண்ணற்ற இனத் தொகுதி — Infinite population
எதிரிடை இயைபு — Negative association
எதிரிடைத் தொடர்பு — Negative correlation
எதிரிடைப் பண்பு — Negative attribute
எல்லை — Limit

ஏ

ஏறுகின்ற — Ascending
ஏறுவரிசை — Ascending order, Ascending order of magnitude

ஒ

ஓகைவ் — Ogive
ஒட்டுறவு — Correlation
ஒப்பீட்டு நிகழ்வெண் — Relative frequency
ஒருபடித்தான — Homogeneous
ஒரு பரவலின் நண்பாதி இடைவெளி — Probable error
ஒருபடித்தான சோதனை — Test of homogeneity
ஒருவால்புறச் சோதனை — One tail test
ஒன்றையொன்று விலக்கும் தன்மைத்தான நிகழ்ச்சி — Mutually exclusive events
ஒற்றைப்படை யெண் — Odd number

க

கட்டுமானம் — Constraint
கட்டின்மை எண்ணிக்கை — Degrees of freedom
கணக்கிடு — Calculate
கணக்கியல் எதிர்பார்ப்பு — Mathematical expectation
கணக்கியல் நிகழ்திறன் — Mathematical probability
கணிப்பு — Calculation
கணிதம் — Mathematics
கண்டறிதல் — Observation
கண்டறிந்த நிகழ்வெண் — Observed frequency
கருதுகோள் — Hypothesis
கருதுகோள் சோதனை — Testing of hypothesis

கருதுகோள் நிகழ்வெண் — Theoritical frequency
கருதுகோள் முழுமைத் தொகுதி — Hypothetical population

கா

காரண விளைவுத் தொடர்பு — Cause and effect relationship
காரணீயப் பெருக்கம் — Factorial
கால்மானம் — Quartile
கால்மான வீச்சு — Quartile range
கால்மான விலக்கம் — Quartile deviation

கீ

கீழின ஒகைவ் — Less than ogive
கீழின வளர்நிகழ்வெண் — Less than cumulative frequency
கீழின வளர்நிகழ்வெண் வளைகோடு — Less than ogive
கீழ் எல்லை — Lower limit
கீழ்க் கால்மானம் — Lower quartile

கு

குழந்தை இறப்பு விகிதம் — Infant mortality rate
குவிவு — Cumulative
குறி — Sign

கூ

கூட்டுச் சராசரி — Arithmetic mean
கூட்டெண் — Positive number
கூருணர்வுத் திறம் — Sensitiveness
கூறு — Sample
கூறெடுத்தல் — Sampling
கூறெடுத்தல் முறைகள் — Sampling methods

கை

கை-வர்க்கம் — Chi-square
கை-வர்க்கச் சோதனை — Chi-square test
கை-வர்க்கப் பட்டியல் — Chi-square table
கை-வர்க்கப் பரவல் — Chi-square distribution

கோ

கோட்டம்	—	Skew
கோட்ட அளவைகள்	—	Measures of skewness
கோட்டக் கெழு	—	Co-efficient of skewness
கோட்டமுடைப் பரவல்	—	Skewed distribution
கோட்டமுடை வளைகோடு	—	Skewed curve
கோணம் (angle)	—	Degree

ச

சதம்	—	Percentage
சதவீத நிகழ்வெண்	—	Percentage frequency
சம அளவு இடைவெளிப் பிரிவு	—	Uniform class interval
சம அளவில்லா இடைவெளிப் பிரிவு	—	Ununiform class interval
சமச்சீரின்மை	—	Asymmetry
சமச்சீரில்லாப் பரவல்	—	Asymmetrical distribution
சமவாய்ப்புக் கொடுத்தெடுக் கப்பட்ட மாதிரி	—	Random sample
சமவாய்ப்புத் திட்டஅமைப்பு	—	Randomised design
சமவாய்ப்புக் கட்டுதிட்ட அமைப்பு	—	Randomised block design
சமவாய்ப்புடை நிகழ்ச்சிகள்	—	Equally likely events
சமன்பாடு	—	Equation
சராசரி (பொதுப்பெயர்)	—	Average
சரிபார்க்கும் குறி	—	Tally mark
சரிவு	—	Slope

சா

சாதக நிகழ்ச்சி	—	Favourable events
சார்பு	—	Function
சார்புடைய, சார்ந்த	—	Dependent
சார்புடைய மாறி	—	Dependent variable
சார்புடை நிகழ்ச்சி	—	Dependent event
சார்பற்ற நிகழ்ச்சி	—	Independent event
சார்பு தெரிந்த முழுமைத் தொகுதி	—	Known population
சார்பு தெரியாத முழுமைத் தொகுதி	—	Unknown population

சிதறல் — Dispersion, Scatter
சிதறல் அளவை — Measures of dispersion
சிதறல் கெழு — Co efficient of dispersion
சிதறல் விளக்கப் படம் — Scatter diagram
சிறப்பு வரம்பு — Level of significance
சிறப்புக் காணும் சோதனை — Test of significance
சிறிய அளவு மாதிரி — Small sample
சிறுமம் — Minimum

சீ

சீரான பரவல் — Symmetricial distribution
சீரற்ற பரவல் — Asymmetrical distribution
சீர்படா விவரங்கள் — Raw data

செ

செயற்கை உறுப்புகள் கொண்ட முழுமைத் தொகுதி — Artificial population
செய்முறைத் திட்டம் — Design of experiments

சே

சேர்மானம் — Combination

சோ

சோதனை அளவை — Test statistic
சோதனைத் திட்ட அமைப்பு — Design of experiments

த

தரம் — Rank
தளை — Constrain
தற்செயல் தொடர்பு — Casual relationship
தனித்த மாறி — Discrete variable

தி

திட்ட அமைப்பு — Design
திட்ட அலகு — Standard unit
திட்டப் பிழை — Standard error
திட்ட விலக்கம் — Standard deviation

திரள் — Aggregate
திருத்தக் காரணி — Correction factor
திரும்பச் செய்தல் — Replicate
திறமுடைய மதிப்பீடு — Efficient estimation

தொ

தொகுத்த செய்திகள் — Data
தொடர்ச்சியான — Continuous
தொடர்ச்சியான பரவல் — Continuous distribution
தொடர்ச்சியற்ற பரவல் — Discrete distribution
தொடர்ச்சியற்ற மாறி — Discrete variable
தொடர் மாறி — Continuous variable
தொடர்பு சமன்பாடு — Regression equation
தொடர்பற்றநிலைச் சோதனை — Test of independence

ந

நம்பக இடைவெளி — Confidence interval
நம்பக எல்லை — Confidence limit
நம்பக வரம்பு — Confidence limit

நி

நிகழ்திறன் — Probability
நிகழ்திற அடர்த்திச் சார்பு — Probability density function
நிகழ்திறப் பிழை — Probable error
நிபந்தனை நிகழ்திறன் — Conditional probability
நிகழ்வெண் — Frequency
நிகழ்வெண் சார்பு — Frequency function
நிகழ்வெண் செவ்வகப்படம் — Histogram
நிகழ்வெண் பட்டியல் — Frequency table
நிகழ்வெண் பரவல் — Frequency distribution
நிகழ்வெண் பலகோணம் — Frequency polygon
நிகழ்வெண் வளைகோடு — Frequency curve
நிகழ்ச்சி — Event
நிமிர்நிலை வரிசை — Column
நிலையான முழுமைத் தொகுதி — Static population

நே

நேர் இயைபு — Positive association
நேர் எண் — Positive number

நேர்கோட்டம்	—	Positive skewness
நேர்கோட்டுக்குரிய	—	Linear
நேர் தொடர்பு	—	Positive correlation

ப

பகடை	—	Die
பகுதி	—	Partial
பகுதி இயைபு	—	Partial association
பகுதித் தொடர்பு	—	Partial correlation
பகுப்பாய்வு	—	Analysis
பகுப்புமுறை இயைபிலா மாதிரி	—	Stratified random sampling
பட்டியல்	—	Table
பண்பு	—	Attribute
பண்பின் அடிப்படையான விவரங்கள்	—	Qualitative data
பண்புகளின் இயைபு	—	Association of attributes
பல்படித்தான	—	Heterogeneous
பயன்முறைப் புள்ளியியல்	—	Applied statistics
பரவல்	—	Distribution
பருமன்	—	Size, magnitude

பா

பாகை	—	Degree
பாய்ஸான் பரவல்	—	Poisson distribution

பி

பிரிவு	—	Class
பிரிவு இடைவெளி	—	Class interval
பிரிவுக் கீழ்எல்லை	—	Lower limit of the class interval
பிரிவு நிகழ்வெண்	—	Class frequency
பிரிவு மேல்எல்லை	—	Upper limit of the class interval
பிரிவெண்	—	Class mark, mid x
பிரித்தமைத்தல்	—	Classification
பிளவு	—	Gap
பிறழ்ச்சியற்ற மதிப்பீடு	—	Unbiased estimate
பிறழ்ச்சியான மதிப்பீடு	—	Biased estimate
பிறப்பு விகிதம்	—	Birth rate

பு

புள்ளியியல் (as data) — Statistics (collection facts)
புள்ளியியல் (as science) — Statistics (statistical methods)
புள்ளியியல் அளவை — Statistic
புள்ளியியல் அனுமானம் — Statistical inference
புள்ளியியல் நிகழ்திறன் — Statistical probability

பூ

பூரணத் தன்மைத்தான நிகழ்ச்சி — Exhaustive events

பெ

பெருக்கற் சராசரி — Geometric mean
பெருமம் — Maximum
பொதுவிதி — General law
பொதுவிதி காண் — Generalise
பொருத்தச் செம்மை — Goodness of fit
பொருளில்லாத் தொடர்பு — Nonsence correlation

போ

போலி இயைபு — Spurious association
போலித் தொடர்பு — Spurious correlation
போதுமான மதிப்பீடு — Sufficient estimate

ம

மக்கட் கணிப்பு, முழுமைத் தொகுதிக் கணிப்பு — Census
மக்கள் மதிப்பீடு — Population —census
மணி வடிவமான — Bell shaped
மதிப்பிடம் — Rank
மதிப்பிடு — Estimate
மதிப்பீடு — Estimation
மதிப்புத் தெரியாத — Unknown
மதிப்பிடத் தொடர்பு — Rank correlation
மதிப்பிடத் தொடர்பு கெழு — Rank correlation co-efficient
மறுக்கத்தகு கருதுகோள் — Null hypothesis
மறுப்பு மண்டிலம் — Rejection region

மா

மாதிரி	—	Sample
மாதிரி அளவு	—	Sample size
மாதிரி அளவையின் பரவல்	—	Sampling distribution
மாதிரி ஏற்ற இறக்கம்	—	Sampling fluctuation
மாறி	—	Variable
மாறிலி	—	Constant
மாறிகளின் தொடர்பு	—	Regression
மாறிகளின் தொடர்பு கெழு	—	Regression co-efficient
மாறுநிலை மண்டிலம்	—	Rejection region
மாறுபாடு	—	Variation
மாறுபாட்டளவை	—	Variance
மாறுபாட்டளவைகள்	—	Measures of variation
மாறுபாட்டுக் கெழு	—	Co-efficient of variation
மாறுபாட்டளவை ஆய்வு	—	Analysis of variance

மி

மிகப் பொருத்தமான	—	Best fit

மு

முகடு	—	Mode
முடிவுள்ள	—	Finite
முதல் கால்மானம்	—	First quartile
முதலடுக்குத் தளை	—	Linear constraint
முதலடுக்குக் கட்டுமானம்	—	Linear constraint
முழுமை	—	Universe
முழுமைத் தொகுதி	—	Population
முழுமைத் தொகுதி அளவை	—	Parameter
முழுமைத் தொகுதிக் கணக்கீடு	—	Census
முரண்படாத மதிப்பீடு	—	Consistent estimate

மூ

மூவழிப் பாகுபாடு	—	Three way classification
மூன்றாவது கால்மானம்	—	Third quartile

மே

மேலின ஓசைவ்	—	Greater than ogive
மேலின வளர்நிகழ்வெண்	—	Greater than cumulative frequency

மேலின வளர்நிகழ்வெண் வளைகோடு — Greater than cumulative frequency curve
மேல் எல்லை — Upper limit
மேல் கால்மானம் — Upper quartile

மை

மையநிலைப் போக்கு — Central tendency
மையநிலைப் போக்கு அளவைகள் — Measures of central tendency
மையப் புள்ளி — Mid x

ரா

ராண்டம் — Random
ராண்டம் எண்கள் — Random numbers
ராண்டம் ஏற்ற இறக்கம் — Random fluctuation
ராண்டம் மாதிரி — Random sample
ராண்டம் மாறி — Random variable

லே

லேடின் சதுரம் — Latin square
லேடின் சதுரத் திட்ட அமைப்பு — Latin square design

வ

வரம்புடை முழுமைத் தொகுதி — Finite population
வரிசை — Arrays
வரிசைமாற்றம் — Permutation
வரையுற்ற — Finite
வரையுற்ற முழுமைத் தொகுதி — Finite population
வரைபடம் — Graph
வரைபட அமைப்பு — Graphic presentation
வளர் — Cumulative
வளர்நிகழ்வெண் — Cumulative frequency
வளர்நிகழ்வெண் வளைகோடு — Cumulative frequency curve

வா

வாய்பாடு	— Formula
வாழ்க்கைப் பட்டியல்	— Life table

வி

விகித சமம்	— Proportion
விலக்கம்	— Deviation

வீ

வீச்சு	— Range